

Tartu Ülikool  
Humanitaarteaduste ja kunstide valdkond  
Ajaloo ja arheoloogia instituut  
Infokorralduse õppekava

**Lilian Neerut**

**TEADUSANDMETE HALDAMISE JA JAGAMISE PRAKTIKAD  
TARTU ÜLIKOOLI LOODUS- JA TÄPPISTEADUSTE NING  
SOTSIAALTEADUSTE VALDKONNA TEADLASTE JA  
DOKTORANTIDE SEAS**

**Magistritöö**

Juhendajad: Kurmo Konsa, PhD

Mai Beilmann, MA

Tartu 2016

# SISUKORD

SISSEJUHATUS .....	3
1 TEADUSANDMED JA AVATUD TEADUS .....	5
1.1 Teadusandmete määratlus .....	5
1.2 Uued võimalused teaduskommunikatsioonis.....	6
1.3 Avatud teaduse kontseptsioon .....	7
1.4 Teadusandmete jagamise olulisus.....	9
1.5 Andmed, mida ei saa jagada .....	10
1.6 Andmete uputus ja kadu .....	10
1.7 Konkreetsete poliitikate ja juhiste vajadus .....	11
1.8 Teadlaste kahtlused teadusandmete jagamisel.....	13
1.9 Huvipoolte rollid ja kohustused.....	14
1.10 Raamatukogu roll avatud teaduse toetamisel.....	15
2 TEADUSANDMETE HALDAMINE .....	17
2.1 Teadusandmete elutsükkel.....	17
2.2 Andmehaldusplaan.....	18
2.3 Andmerepositooriumid .....	19
3 UURINGU LÄBIVIIMINE .....	20
3.1 Uuringu eesmärk.....	20
3.2 Metoodika .....	20
3.3 Valim .....	21
4 TULEMUSTE ANALÜÜS .....	23
4.1 Teadusandmete kogumine .....	23
4.2 Teadusandmete arhiveerimine .....	26
4.3 Teadusandmete kättesaadavus, nähtavus ja taaskasutatavus .....	30
4.4 Infrastruktuur ja teenused .....	38
5 ARUTELU .....	41
KOKKUVÕTE .....	46
SUMMARY .....	49
KASUTATUD KIRJANDUS .....	50
LISA 1 .....	59
LISA 2 .....	61
LISA 3 .....	62

## SISSEJUHATUS

Tänu teaduse ja tehnika arengule on teadustöö käigus kogutavate ning loodavate andmete hulk varasemaga võrreldes tohutult kasvanud ning kasvab ilmselt veelgi. Tänapäeva tehnoloogia soodustab rahvusvahelist teadlastevahelist koostööd, võimaldades kiiresti liigutada suuri andmehulki ning neid omavahel kombineerida ja analüüsida, muutes seeläbi teaduse peamise ülesande, lahendada probleeme, efektiivsemaks. Teisalt on see kaasa toonud olukorra, kus tegelikult puudub ülevaade kõikidest kogutud andmetest. Väga palju andmeid läheb kaduma, sest neid ei säilitata korralikult vähese oskuste või motivatsioonipuuduse tõttu.

Samas liigub teadus avatud mõtteviisi suunas, mille järgi avaliku raha eest rahastatud teadusuuringute tulemused (sh teadusandmed) peaksid olema vabalt kättesaadavad ja taaskasutatavad.

Kõik need protsessid on kaasa toonud vajaduse reguleerida andmete säilitamise ja jagamise põhimõtteid. Viimase kümne aasta jooksul on pööratud üha rohkem tähelepanu teadusandmete haldamisele.

Jõuliselt tõusis teema esile 2000ndate lõpus seoses majanduskriisiga ja sellest tulenevalt ka teadusrahastuse vähenemisega. Piiratud finantsvahendite juures tahavad rahastajad olla kindlad, et teadlased kasutavad ressursse parimal viisil ja andmete kadu vältides (Briney, 2015:2). Tänapäevaks on paljud teadusrahastajad, näiteks kõik Suurbritannia teadusnõukogud (Overview of funders..., 2016), viinud granditaotluse juures sisse andmehaldusplaani esitamise nõude. Ka Euroopa Liidu teadusuuringute ja innovatsiooni raamprogrammis Horisont 2020 rakendatakse pilootprojekt avatud juurdepääsuks teadusandmetele ning uurimistoetuse taotlejal on kohustus esitada andmehaldusplaan, kus näidatakse, millised andmed on avaandmed (European Commission, 2016b).

Samuti nõuavad mitmed teadusajakirjad, näiteks PLOS (PLOS, 2016), teadusartikli juures alusandmete avaldamist.

Kõige olulisem on teadusandmete haldamine teadlasele endale. Reeglite kohaselt hallatud andmed on talle endale kergesti leitavad ja võimaldavad tõestada oma teadustulemusi. Avaldatud andmeid on võimalik tsiteerida ning seeläbi suureneb teadustöö nähtavus. Samuti suurenevad koostöövõimalused teiste teadlastega.

Magistritöö teoreetilise osa eesmärk on anda ülevaade teadusandmete säilitamise, kättesaadavaks tegemise ja taaskasutamise vajalikkusest ning problemaatikast.

Magistritöö empiirilises osas uuritakse, kas ja kuidas koguvad, jagavad, säilitavad ja taaskasutavad teadusandmeid Tartu Ülikooli loodus- ja täppisteaduste valdkonna ja sotsiaalteaduste valdkonna teadlased ja doktorandid ning millisest lisainformatsioonist ollakse huvitatud.

Kogu maailmas on ülikoolide raamatukogud üha enam seotud avatud teaduse toetamisega. Suurt tähelepanu pööratakse teadusandmete haldamise teenuste väljatöötamisele. Esimesena Eestis pakub ka Tartu Ülikooli raamatukogu teadusandmete arhiveerimisvõimalust institutsionaalses repositooriumis (<http://datadoi.ut.ee/>) ning teadusandmete haldamisega seotud nõustamist ning koolitusi.

Teemavalik on otseselt seotud autori tööalase huviga. Tartu Ülikooli raamatukogu teadusandmete spetsialistina on oluline olla kursis sellega, kuidas teadlased teadusandmeid haldavad ning millist lisateavet vajavad. Uuringu tulemused on abiks ülikooli raamatukogu teadusandmete haldamise teenuste väljaarendamisel.

Eestis on varasemalt teadlaste andmehaldusalaseid arvamusi uurinud Eesti Teadusagentuur 2014. aastal (Muuli, 2014).

Töö koosneb viiest peatükist. Esimene peatükk keskendub teadusandmete säilitamise, jagamise ja taaskasutamise olulisusele ning võimalikele takistustele. Samuti tuuakse välja huvipoolte rollid ja kohustused ning avatud teaduse kontseptsioon. Teine peatükk selgitab teadusandmete haldamise põhimõtteid. Kolmandas peatükis sõnastatakse uurimisteema ja -küsimused ning kirjeldatakse uuringu metoodikat ja valimit. Neljandas peatükis esitatakse tulemuste analüüs ning viiendas tulemuste arutelu.

# 1 TEADUSANDMED JA AVATUD TEADUS

## 1.1 Teadusandmete määratlus

Teadusandmete kirjeldamiseks puudub kõikehõlmav definitsioon. Kõige üldiselt öeldes on teadusandmeteks kõik andmed, mis on kogutud või loodud teadustöö käigus. Ühelt poolt on teadusandmed teadustöö tulemus, aga samas ka selle allikaks.

Teadusandmed võivad sõltuvalt valdkonnast andmetüübilt ja formaadilt esineda väga erineval kujul. Näiteks tekstid, pildid, helifailid, laboripäevikud, algoritmid, transkriptsioonid, esitluste slaidid jne (EDINA, 2015).

Teadusandmed võivad olla näiteks vaatlusandmed (*observational data*), eksperimentaalandmed (*experimental data*), simulatsioonide andmed (*simulation data*) ja kokkuvõetud andmed (*compiled data*) (Briney, 2015: 6). Vaatlusandmed saadakse tavaliselt kindlal ajal ja kohas konkreetse juhtumi uurimisel. Eksperimentaalandmed tekivad kontrollitud keskkondades katseid tehes ja neid korrates. Simulatsioonide andmeid produtseeritakse arvutisimulatsioonidega ning kokkuvõetud andmed tekivad siis, kui muudest allikatest pärinevad andmed kogutakse kokku teistkordseks kasutamiseks, näiteks metaanalüüsi läbiviimiseks (Briney, 2015: 6).

Teadusandmed võivad olla sekundaarandmed ja primaarandmed. Sekundaarandmed on kogutud teiste teadlaste poolt ja primaarandmed teadlase enda poolt uurimisprojekti raames (Herzog, 2015:6).

Üks võimalus teadusandmete kirjeldamiseks on: „Teadusandmed on andmed, mida kasutatakse esmase allikana toetamaks tehnilisi või teaduslikke uuringuid, teadustööd, õpinguid või kunstialast tegevust ja mida kasutatakse teadustöö käigus tõendusmaterjalina ja/või mida teaduskogukond peab teadustöö leidude ning tulemuste paikapidavuse tõendamiseks üldiselt vajalikuks. Igasugune muu digitaalne või ka mittedigitaalne sisu võib potentsiaalselt saada teadusandmeteks. Teadusandmed võivad hõlmata eksperimentaal-andmeid, vaatlusandmeid, ekspluatatsioonandmeid, kolmandatelt osapooltelt saadud andmeid, avaliku sektori andmeid, seireandmeid, töödeldud andmeid või uue otstarbe saanud andmeid“ (CASRAI dictionary, 2015)

## 1.2 Uued võimalused teaduskommunikatsioonis

Läbi ajaloo on teaduse tegemise viisid vägagi suurel määral muutunud. See on ühelt poolt seotud teaduse enda arenguga, aga teisalt ka ühiskonna ja tehnoloogia pideva muutumisega.

Koos infotehnoloogia arenguga on teaduses toimunud fundamentaalsed muutused. Muutunud on nii andmete hankimine kui ka nende töötlemine ja kasutamine.

Piltlikult öeldes ei istu teadlased enam teleskoobi taga, vaid varasemad teleskoobid on asendunud komplekssete instrumentidega, mis edastavad andmed otse andmekeskustesse, kust teadlane vaatlusandmed kätte saab (Gray, 2009: xix).

Andmehulkade kasvu ja tehnoloogia kiire arengu valguses on toimumas nihe ka teaduskommunikatsioonis. Traditsiooniliselt toimub teadustulemuste esitamine teistele teadlastele läbi eelretsenseerimise protsessi, mis aitab uuringutulemusi enne teadusartikli avaldamist kontrollida (Committee on Ensuring..., 2009: 3).

Arvatakse, et kuigi eelretsenseerimine jääb peamiseks teadustulemuste kehtivuse üle otsustamise praktikaks, on seda keeruline kasutada digitaalsete andmete kogumise, analüüsimise ja interpreteerimise asjakohasuse hindamisel. Õnneks võimaldab digitaalsete tehnoloogiate areng kindlustada teadusandmete terviklikkuse (*data integrity*) läbi suurema avatuse ja läbipaistvuse (Committee on Ensuring..., 2009: 3). Teadusandmed on terviklikud, kui nad on kontrollitud, esitatud täielikult ja muutmata kujul (Committee on Ensuring..., 2009: 22).

Tänaseks on näiteks Cornelli ülikoolis välja töötatud automaatne teadusandmete kvaliteedi kontrolli ja tulemuste kordamise teenus (*Data Quality Review and Reproduction of Results Service*) (Arguillas & Block, 2016).

Olulise muutusena teaduskommunikatsiooni protsessis on välja toodud ka võimalused teadlastevaheliseks tihedamaks suhtlemiseks. Uued infrastruktuurid võimaldavad varasemast palju suuremate andmehulkade jagamist ning see on nihutanud paradigma uute, jagamisel ja koostööl põhinevate teadustöömetodoloogiateni (Paterson, Lindsay, Monotti, & Chin, 2007: 116). Suber (2012) nimetab info kiiresti ja tasuta jagamise võimalust internetis „juurdepääsu revolutsiooniks“ (*access revolution*) (Suber, 2012: 1)

### 1.3 Avatud teaduse kontseptsioon

Teadusmaailm liigub üha enam avatud mõtteviisi suunas. Avatud teadus (*open science*) juhindub kuuest põhiprintsiibist: avatud juurdepääs (*open access*), avaandmed (*open data*), avatud lähtekood (*open source*), avatud metodoloogia (*open methodology*) (Kraker, Leony, Reinhardt, & Beham, 2011: 645) ning avatud eelretsenseerimine (*open peer review*) ja avatud hariduse ressursid (*open educational resources*) (Kasberger, 2013).

- Avatud juurdepääs tähendab teadustulemustele vaba ja piiranguteta juurdepääsu internetis. Avatud juurdepääs on avatud teaduse keskmeks (Kraker et al., 2011: 645-646).
- Avaandmed on teadustöö käigus kogutud andmed, mis tehakse internetis avatud juurdepääsuga kättesaadavaks (Murray-Rust, 2008).
- Avatud lähtekood lubab olemasolevat tarkvaralist lähtekoodi vabalt kasutada (Kraker et al., 2011: 646).
- Avatud metodoloogia tähendab detailset kirjeldust uurimuse läbiviimise kohta, s.h kuidas andmeid koguti, kirjeldati ja analüüsiti (Kraker et al., 2011: 647).
- Avatud eelretsenseerimine tähendab teadustulemuste kvaliteedi kinnitamist läbipaistva ja jälgitava avatud eelretsenseerimisprotsessi kaudu (Kasberger, 2013).
- Avatud hariduse ressursside all mõeldakse vabalt kättesaadavaid ja kasutatavaid õppematerjale (Kasberger, 2013).

Teadusandmete kättesaadavaks tegemise ja avatud teaduse kontseptsiooni toetavad ning aitavad teadvustada mitmed rahvusvahelised kokkulepped ja soovitusel.

Teadusandmete jagamise ja rahvusvahelise koostöö teadvustamise lepingulised alged ulatuvad juba 1959. aastasse, mil Washingtonis sõmiti 12 riigi vahel Antarktika leping. Muuhulgas sätestab leping, et: „Antarktika teadusvaatlusi ja nende tulemusi vahetatakse ja tehakse vabalt kättesaadavaks“ (The Antarctic Treaty, 1961).

1996. aastal sõnastasid teadlased poolt „Human Genome Project“ raames nn Bermuda printsiibid inimgeeni sekventeerimisandmete säilitamise ja avalikult kättesaadavaks tegemise kohta (Summary of the..., 1997).

USA National Research Council (1997) poolt tellitud uuringus märgitakse: „Andmete väärtus seisneb nende kasutuses. Avaliku sektori rahastatud teadusuuringute tulemusena saadud teadusandmete vahetamise edendamiseks peaks rahvusvaheliseks normiks saama teadusandmetele täieliku ja avatud juurdepääsu tagamine“ (Committee on Issues..., 1997: 10).

Avatud juurdepääsu ideoloogia edendamine hoogustus 21. sajandi esimesel kümnendil, mil võeti vastu ridamisi deklaratsioone ja soovitusi. Järgnevalt mõned näited.

Olulisteks teetähisteks teadusinformatsiooni ja kultuuripärandi säilitatamise ja internetis avatud juurdepääsu võimaldamise mõtteviisi teadvustamisel on 2002. aastal kirja pandud „Budapest Open Access Initiative“ (Budapest Open Access Initiative, 2002) ning 2003. aastal sõnastatud „Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities“ (Berlin Declaration on..., 2003).

Berliini deklaratsioonile järgnenud 2004. aastal allkirjastatud dokumendi „OECD Principles and Guidelines for Access to Research Data from Public Funding“ soovituste kohaselt peaksid avaliku raha toetusel kogutud ja loodud teadusandmed olema teaduspotentsiaali suurendamiseks vabalt kättesaadavad ning taaskasutatavad (OECD Principles and..., 2007).

Teadusandmete ekspertide grupi poolt Euroopa Komisjonile koostatud raportis keskendutakse vajadusele arendada välja e-infrastruktuur, mis toetab teadusandmete globaalset kättesaadavust, usaldusväarsust, kasutamist ja taaskasutamist. „Füüsiline ja tehniline infrastruktuur muutub nähtamatuks ning andmetest endist saab infrastruktuur – väärtuslik vara, millele tuginedes arenevad nii teadus, tehnoloogia, majandus kui ka ühiskond“ (The High Level..., 2010: 24).

2012. aastal allkirjastas Euroopa Komisjoni asepresident Neelie Kroes soovitused teadusinfo kättesaadavuse ja säilitamise kohta (Euroopa Komisjon, 2012).

Koos digitaalse ühisturu strateegia avalikustamisega 2015. aasta mais tutvustas Euroopa Komisjon algatust, mille eesmärk on luua Euroopa avatud teaduse pilv (*European Open Science Cloud*). Avatud teaduse pilv peaks pakkuma kasutajatele virtuaalses keskkonnas tasuta, avatud ja mugavat teenust teadusandmete majutamiseks, haldamiseks, analüüsimiseks ja taaskasutamiseks nii, et kõik teadusandmed oleksid lingitud teadlase teiste teadustegevustega (European Commission, 2016a).

Eesti teadus- ja arendustegevuse ning innovatsiooni strateegia 2014-2020 „Teadmistepõhine Eesti“ sätestab, et tuleb „soodustada avatud juurdepääsu (ingl *open access*) avaliku rahastuse eest saadud uurimistöö tulemustele ja teadusandmetele“ (Eesti teadus- ja arendustegevuse..., 2014: 16).



## 1.4 Teadusandmete jagamise olulisus

Teadustulemuste esitlemine ja jagamine on teadlaste jaoks ülioluline. Traditsiooniliselt on seda tehtud artikleid teadusajakirjades publitseerides või kirjutades monograafiaid (Hallik, 2014). Tänapäeva tehnoloogiline võimekus ühendada teadusandmed teaduskirjandusega ja vastupidi ning nende vastastikune koostoimevõime tõstab „teadusinformatsiooni kiirust“ (*information velocity*) ja aitab seeläbi kaasa teadlaste tulemuslikkusele (Gray, 2009: xxv).

Van den Eynden jt (Van den Eynden, Corti, Woollard, Bishop, & Horton, 2011: 3) rõhutavad teadusandmete jagamise olulisuse juures, et see soodustab innovatsiooni, aitab kaasa teaduse arengule, suurendab teadlastevahelist koostööd, võimaldab uuringu tulemusi põhjalikult uurida. Samuti suurendab uuringu läbipaistvust, aitab kaasa uurimismeetodite arengule ja kontrollitavusele. Teadusandmete jagamine vähendab andmete kogumise dubleerimist ja hoiab seeläbi kulusid kokku. Läbi tsiteeringute suureneb uuringu nähtavus ja mõjukus. Ka Piwowar ja Vision märgivad, et teadusartiklid, millele on lisatud uuringu andmed saavad rohkem tsiteeringuid, kui sarnase temaga artiklid, mille juures andmed puuduvad (Piwowar & Vision, 2013: 17).

Selle kõrval ei ole sugugi väheoluline ühiskonna vajadus juurdepääsuks teadusandmetele. Äriettevõtted kasutavad uusi teadmisi toodete ja rakenduste arendamiseks. Valitsuste otsused peaksid tuginema andmetel põhinevatel analüüsidel. Kodanike juurdepääs teadusandmetele aitab kaasa kollektiivse teadusliku teadmuse (*collective scientific knowledge*) arengule (Corti, Van den Eynden, Bishop, & Woollard, 2014: 1).

Teadusandmete taaskasutamine, eriti, kui on võimalik kombineerida erinevaid andmestikke, aitab kaasa uutele avastustele (Piwowar & Vision, 2013: 1)

Tulevastele teadlastele mõeldes tuleb silmas pidada, et teadusandmed on heaks baasiks õpetamisel. Samuti räägitakse sellest, et teadusandmete kättesaadavaks tegemine peaks olema üheks kriteeriumiks teadlase teadusväljundi hindamisel (Van den Eynden et al., 2011: 3).

Teadusandmete jagamise üks eesmärkidest on, et andmeid oleks võimalik taaskasutada. Zimmerman nimetab andmete taaskasutamiseks seda, kui konkreetse uuringu tarbeks kogutud andmeid saab kasutada millegi täiesti uue uurimiseks (Zimmerman, 2008: 634).

Millal ja kellele võivad kogutud andmed huvi pakkuda on raske ennustada. Nii toob Bell näite, kuidas planeetide liikumise seaduse avastas Kepler tänu Tycho Brache süstematiseeritud astronoomiliste vaatluste kataloogile (Bell, 2009: xi).

Washingtoni ülikooli geoloogiaprofessori Brian Atwater'i juhtimisel uuriti võimalikke maavärinaohtusid Ameerikas Vaikse ookeani loodekalda piirkonnas. Uuringus kasutati ka 1960ndatel ja 1970ndatel kogutud doktoritöodes avaldatud sedimentoloogialaseid andmeid, mis ei olnud algselt maavärinauuringute jaoks kogutud (Atwater, Carson, Griggs, Johnson, & Salmi, 2014).

USA klimatoloogid on algatanud projekti „Old Weather: Whaling“, mille raames transkribeeritakse 19. sajandi vaalapüüdjate logiraamatuid. Ajaloolisi andmeid jää ja jäämägede kohta plaanitakse kasutada kliimamuutuste uuringutes (Old Weather: Whaling, 2016).

Smith ja Hatemi uuring üksikisiku geneetilise baasi ja tema poliitiliste vaadete seoste kohta on hea näide katsest sünteesida erinevate valdkondade andmeid ja metoodikaid (Smith & Hatemi, 2012).

## **1.5 Andmed, mida ei saa jagada**

Vaatamata eelnevalt väljatoodud argumentidele teadusandmete jagamise olulisusest, on hulgaliselt andmeid, mida ei saa mitmetel põhjustel avalikult kättesaadavaks teha. Tegemist võib olla näiteks tundlike andmetega, ärihuvidega või riigisaladusega seotud andmetega (Committee on Ensuring..., 2009 :5).

Eriti hoolikalt tuleb ringi käia delikaatsete isikuandmetega või mistahes lisainformatsiooniga andmestikus, mis võimaldab isikut identifitseerida (Briney, 2015: 25). Eestis on delikaatsed isikuandmed välja toodud isikuandmete kaitse seaduses (Isikuandmete kaitse seadus, 2007).

Kui vähegi võimalik, siis soovitatakse tundlikud andmed anonümiseerida, eemaldades andmestikust isikut tuvastada võimaldava otsese ja kaudse informatsiooni (Briney, 2015: 108).

AS Cybernetica meeskond on välja arendanud isikuandmete ja ärisaladuste turvalise analüüsi ja jagamise süsteemi Sharemind (Cybernetica AS, 2016).

## **1.6 Andmete uputus ja kadu**

Tänu tehnoloogia arengule on viimase viieteistkümne aasta jooksul kasvanud digitaalse informatsiooni hulk eksponentsiaalselt ning see trend on jätkuv (HRK German Rectors..., 2014: 2).

21. sajandi teadusandmete kasvu on nimetatud ka andmete uputuseks (*data deluge*) (Hey & Trefethen, 2003: 809). Gray räägib andmete jäämäest (*data iceberg*), mõeldes selle all suurt hulka teadusandmeid, mida kogutakse, aga mis tihtipeale jäävad pärast teadusartikli avaldamist täielikult kasutamata, süstematiseerimata ja hooldamata (*curated*) (Gray, 2009: xvii).

Bell märgib, et 20. sajandil põhinesid teaduslikud teooriad andmetel, mis olid tihtipeale fikseeritud ainult teadlaste isiklikes teadusmärkmetes. Eriti just väikelaborite andmed ei ole hiljem enamasti kättesaadavad (Bell, 2009: xii).

Teadusandmete võimalikus kaos võib olla mitmeid põhjuseid. Näiteks võib arvuti kõvaketas töötamast lakata. Tihti on digitaalsed andmed talletatud tehnoloogiliselt kiiresti vananevates formaatides, mida tänapäeva arvutiprogrammid ei suuda lugeda (Committee on Ensuring..., 2009: 96). Andmekao põhjuseks võib olla ka näiteks arvutiviirus (Briney, 2015: 116).

Teadlane võib vahetada institutsiooni või pensioneeruda (Committee on Ensuring..., 2009:96) ning, kui andmed on korrastamata ja kirjeldamata, siis ei ole need enam kasutatavad. Andmed võivad kaduda ka arvuti varguse, inimliku unustamise või isegi looduskatastroofi tõttu (Briney, 2015:116).

Hulgaliselt teadusandmeid kaob eelkõige puudulike säilitamistingimuste ja -oskuste tõttu. Väärtuslike andmete kadu on peamiselt väiksemate ja üksikuurijate probleem, sest neil ei pruugi olla piisavalt oskusi ja rahalisi vahendeid andmete arhiveerimiseks ettevalmistamiseks. Suuremates projektides on tavaliselt andmete salvestamise ja pikaajalise säilitamise eest hoolitsetud. Näiteks molekulaarbioloogia- ja ökoloogiauuringud on tänapäeval suurte andmetihedate teadusprojektide osad. Sotsiaalteadustes on andmete pikaajaliseks säilitamiseks aastakümnete pikkused traditsioonid. (Committee on Ensuring..., 2009: 96-98).

## **1.7 Konkreetsete poliitikate ja juhiste vajadus**

Van de Sompel ja Lagoze meenutavad Gray mõtet, et ka kõige keerulisema tehnilise ülesande kõrval on palju fundamentaalsem probleem, kuidas integreerida küberinfrastruktuuri võimalused inimeste töövoogudesse ja praktikatesse (Van de Sompel & Lagoze, 2009: 198). Ilma konkreetsete rahvusvaheliste ja rahvuslike poliitikateta ja toetava õigusliku raamistikuta jooksevad ideed teadusandmete jagamisest ja koostööst liiva (Fitzgerald, Fitzgerald, & Pappalardo, 2009: 204).

Selleks, et võimalikult palju teadusandmeid säilitataks ning neid oleks võimalik taaskasutada, tuleb kokku leppida reeglid. Avatud teaduse osas kirjeldatud rahvusvaheliste kokkulepete rakendamiseks on vaja kehtestada üksteist täiendavaid poliitikaid rahvusvahelisel, rahvuslikul ning institutsionaalsel tasemel. Poliitikate abil loodetakse kiiremini muuta tavapäraseid harjumusi ja kinnistada uus mõtteviis.

Ka Hallik ja Tarkpea toonitavad, et hetkel kaaluvad hirmud eeldatava kasu üles ning kõik teadlased ei ole veel valmis andmeid vabatahtlikult jagama. Seepärast on kultuuri muutuseks vaja väliseid mõjutajaid (Hallik & Tarkpea, 2016: 82).

Arvestada tuleb sellega, et andmete säilitamine nõuab finantsressursse. Kuigi andmete taaskasutamine muutub aina olulisemaks, ei ole võimalik üha kasvavate andmehulkade juures kõiki andmeid säilitada. Seetõttu on vaja välja töötada konkreetseid pikaajalise säilitamise juhendid teadlastele, teadusasutustele ja teadusrahastajatele (Committee on Ensuring..., 2009: 98).

Uurinud perioodil 1991-2011 ilmunud 516 artikli puhul andmete hilisemat kättesaadavust, leiavad Vines jt, et teadusartiklite aluseks olevate andmekogude kättesaadavus väheneb iga aastaga 17% ning soovivad rahastajatel kiiresti andmete pikaajalise säilitamise huvides vastavad teemat reguleerivad poliitikad kehtestada (Vines et al., 2014: 94).

Toimivate poliitikate kujundamisel on oluline arvestada kõigi huvipoolte arvamusega. Kaks aastat väldanud RECODE (*Policy Recommendation for Open Access to Research Data in Europe*) projekti raames toodi kokku erinevad huvipooled (rahastajad (*funders*), teadusasutused (*research institutions*), andmehaldajad (*data managers*), kirjastused (*publishers*)), et moodustada koostöövõrgustik ning töötada üheskoos välja poliitika avatud juurdepääsu eesmärkideni liikumiseks. Projekti lõppedes sõnastati 10 üldist poliitikasoovitust Euroopa Liidu liikmesriikidele (Recode Project Consortium, 2014):

1. Töötada välja teaduse avaandmete juurdepääsu üldine reguleeritud poliitika.
2. Kindlustada teaduse avaandmete juurdepääsuks asjakohase rahastamissüsteemi olemasolu.
3. Töötada välja põhimõtted ja algatused, mille abil premeerida ja innustada teadlasi võimaldama avatud juurdepääsu kõrgekvaliteedilistele andmetele.
4. Selgitada välja tähtsaimad huvipooled ja asjassepuutuvad võrgustikud ning edendada koostööd teaduse avaandmete jätkusuutliku ökosüsteemi loomiseks.
5. Koostada tegevuskava avaandmete pikaajaliseks ja jätkusuutlikuks hooldamiseks ja säilitamiseks.

6. Töötada välja üldised, koostööl põhinevad tehnilised ja taristut puudutavad lahendused, mis võimaldavad avatud juurdepääsu kõrgekvaliteedilistele teadusandmetele ja nende pikaajalist säilitamist.
7. Töötada välja teadusandmete tehnilise ja teadusliku kvaliteedi standardid.
8. Nõuda ühtlustatud avatud litsenseerimise tavadest kinnipidamist.
9. Süstemaatiliselt tegelda teaduse avaandmetega seotud õiguslike ja eetiliste küsimustega.
10. Toetada koolituste ja õppekavade täiendamise kaudu üleminekut teaduse avaandmete kättesaadavaks tegemisele ja kasutamisele.

Enamiku riikide jaoks on avatud teaduse teema poliitikate tasemel uus ning arenemisjärgus. Riiklikud strateegiad ja poliitikad on avaldatud näiteks Sloveenias (Government of the..., 2015), Iirimaa (Dempster, 2014), Hispaanias (Anglada et al., 2014), Belgias (Brussels Declaration on..., 2012), Portugalis (University of Minho, 2014), Taanis (Ministry of Higher Education and Science, 2014), Rootsis, (Swedish Research Council, 2015), Austrias (Tonta, Doğan, Al, & Madran, 2015), Norras (Karlstrøm & Aasheim, 2014), Leedus (The Research Council..., 2016), Soomes (Ministry of Education..., 2014). Eestis on SA Eesti Teadusagentuur eestvedamisel kokku kutsustud avatud teaduse ekspertkomisjon, mille ülesanne on „välja töötada ettepanekud avatud teaduse riikliku poliitika põhimõtete kujundamiseks“ (Eesti Teadusagentuur, 2015). 2016. aasta veebruaris avaldati dokumendi „Avatud teadus Eestis“ tööversioon, mille järgi plaanitakse ka Eestis sisse viia granditaotluste esitamisel andmehaldusplaanide koostamise nõue (Eesti Teadusagentuuri avatud..., 2016).

## **1.8 Teadlaste kahtlused teadusandmete jagamisel**

Juba ligi 30 aastat tagasi tehtud uuringus tõdeb Ceci, et teadlased ühelt poolt toetavad andmete jagamise põhimõtet kui edasiviivat jõudu teaduses, kuid teisalt muretsevad, et kõik teadlased ei ole siiski valmis avaliku rahastusega toetatud uuringute andmeid kolleegidega jagama, kartes, et see võib takistada nende endi karjäärivõimalusi (Ceci, 1988: 51-52).

Osa põhjusi andmete mittejagamiseks on otseselt teadusvaldkonna spetsiifikaga seotud, kuid enamik neist on siiski universaalsed. Peamiselt on need seotud teaduse tegemise kultuuri ja teadlase enda hoiakutega (Tenopir et al., 2011).

Kõige märkimisväärssem on individuaalne barjäär (Poline et al., 2012: 5). Kardetakse ideede ülevõtmist ja paremate tõlgendusteni jõudmist teiste teadlaste poolt. Samuti on

põhjuseks ebakindlus oma uurimismeetodite osas ning ei taheta näida ebakompetentne (Poline et al., 2012: 6). Savage ja Vickers toovad välja soovi omada eksklusiivset juurdepääsu andmetele, mida on suure vaevaga kogutud (Savage & Vickers, 2009: 3).

Eetilised ja õiguslikud probleemid puudutavad eelkõige inimestega seotud uuringuid. Enne uuringuid võetud informeeritud nõusolekud ei sisalda tihti teavet andmete laialdase jagamise kohta. Samuti kogutakse andmeid tavaliselt konkreetse probleemi uurimiseks (Poline et al., 2012: 6) Savage ja Vickers toovad samuti välja mure uuritava privaatsuse ja teadlase enda autorsuse pärast (Savage & Vickers, 2009: 3).

Tehnilised takistused puudutavad peamiselt väiksemaid uurimisprojekte. Suurte projektide jaoks on tavaliselt välja arendatud projekti raames spetsiaalsed e-platvormid. Väiksemad projektid jagavad tihti andmeid neid ise või failiedastusprotokolli kaudu saates (Poline et al., 2012: 6). Takistuseks andmete jagamisel võivad olla ka teadlaste ebapiisavad teadmised andmehaldusest (Corti, Van den Eynden, et al., 2014: 13). Vähetähtis ei ole hirm andmete väärkasutuse pärast. Ühe takistusena andmete jagamisel on välja toodud ka asjaolu, et sellega ei arvestata akadeemilises hindamisprotsessis (Schäfer et al., 2011: 75). Samuti arvatakse, et andmete ettevalmistamine jagamiseks on liiga ajamahukas (Savage & Vickers, 2009: 9).

## **1.9 Huvipoolte rollid ja kohustused**

Teadusandmete säilivuse ja kättesaadavuse tagamine, vastava mõtteviisi arendamine ja praktikasse rakendamine nõuab kõigi huvipoolte panustamist. Andmete jagamist ei tohiks võtta ainult kui bürokraatlikku nõuet ning täita seda pealiskaudselt, näiliselt kehtestatud nõudmistele vastates. Näiteks on korralike metaandmetega varustamata andmed repositooriumis kasutud, sest nad ei ole leitavad (Leonelli, Spichtinger, & Prainsack, 2015: 13).

Teaduskogukondadel tuleks koostöös teadlastega, rahastajatega ja teadusajakirjadega välja arendada valdkonnapõhised standardid, pakkuda koolitusvõimalusi, tunnustada avalikult andmete jagamise ja säilitamise poliitikaid (Committee on Ensuring..., 2009: 88-117).

Teadlastel lasub esmane kohustus hoolitseda teadusandmete, uurimismeetodite ja muu seotud informatsiooni süstematiseerimise, säilitamise ja kättesaadavaks tegemise eest viisil, et need oleksid teiste teadlaste poolt kasutatavad, välja arvatud põhjendatud erijuhud, mil andmete avaldamine ei ole võimalik (Committee on Ensuring..., 2009: 87-117). Oluline on aktiivselt osaleda ja kaasa rääkida vajaduste välja selgitamisel ning lahenduste

testimisel, standardite ja juhendite välja töötamisel (Committee on Ensuring..., 2009: 116-117; Jones, Whyte, & Pryor, 2013: 3).

Teadusasutused peaksid välja töötama ja teadlaskonnale vahendama konkreetseid teadusandmete haldamise ja jagamise poliitikad ning tagama jätkusuutliku teenuste ja infrastruktuuri toimimise. Samuti tuleks välja töötada karjääristiimulid teadlastele ning toetada andmehaldusalaste oskuste arendamist (Committee on Ensuring..., 2009: 90-91; Jones et al., 2013: 3). Akadeemilises hindamisprotsessis tuleks tulevikus arvestada traditsiooniliste bibliomeetriliste mõõdikute kõrval ka altmeeriliste tulemustega (Leonelli et al., 2015: 14). Altmeetria tähendab, et andmeid teadustulemuste leviku ja mõju kohta kogutakse sotsiaalmeediast, veebist, uudistest, blogidest jm (Priem, Taraborelli, Groth, & Neylon, 2010).

Rahastajad saavad teadusandmete jagamise kultuuri arendamisele kaasa aidata, kehtestades omad teadusandmete haldamise poliitikad (Expert Advisory Group..., 2014: 20-22).

Teadusajakirjad peaksid nõudma teadusartiklite juurde uuringu alusandmete avaldamist ning nende pikaajalise säilitamise kindlustamist. Ajakirjad saavad kaasa rääkida standardite välja töötamises ja rakendamises (Committee on Ensuring..., 2009: 116-117). Nõudes koos artikliga ka alusandmete avaldamist mängivad teadusajakirjad võtmerolli andmete arhiveerimisel ja teadlaskonnale kättesaadavaks (Piwowar & Chapman, 2008; Poline et al., 2012: 7).

## **1.10 Raamatukogu roll avatud teaduse toetamisel**

Ülikoolide raamatukogudel on oluline roll teadlaste toetamisel teadusandmete haldamisel, säilitamisel ja kättesaadavaks tegemisel (Cox & Pinfield, 2014: 300).

Vajalikud eeldused nagu näiteks pikaajalised kontaktid teaduskondadega, oskused ja kogemused informatsiooni standardiseeritud organiseerimiseks ning metaandmete loomiseks, on raamatukogudel olemas. Raamatukogud on olnud ka peamised avatud juurepääsu mõtteviisi propageerijad. Teadusandmete haldamise teemad oleksid infopädevuse kursuste loomulik osa (Cox & Pinfield, 2014: 300). Raamatukogudel on pikaajalised kogemused erinevates vormides info säilitamisel (Henty & others, 2008: 4).

Lyon (2012) näeb raamatukogu ülesannetena teadusandmete haldamise kontekstis (Lyon, 2012: 129-130):

- aidata kaasa ülikoolisest teadusandmete haldamise alasete vajaduste väljaselgitamisele;

- pakkuda kõigile õppetasemetele ja teadlastele andmehalduse planeerimise nõustamist ning töötada välja vastavad juhendmaterjalid;
- pakkuda tehnilist nõustamist teadusandmete kirjeldamisel ja metaandmete loomisel;
- toetada teadusandmete tsiteerimist, omistades andmekogudele digitaalse objekti identifikaatorid (DOI);
- õpetada teadusandmete haldamist;
- pakkuda nõustamist andmehaldusega seotud õiguslike ja eetilistele küsimuste osas;
- aidata hinnata, milliseid andmeid tuleks säilitada ja milliseid mitte;
- koostöös IT osakonnaga töötada välja andmete säilitamise juhendid;
- kindlustada teadusandmete säilimine ja juurdepääs institutsionaalses repositooriumis;
- koguda bibliomeetrilist infot andmekogude kasutamise kohta.

Ka Tartu Ülikooli raamatukogu arengukava aastateks 2016-2020 näeb ette, et raamatukogu osutab „teenuseid teadusandmete avaldamiseks, registreerimiseks ja pikaajaliseks säilitamiseks, liidestamiseks ja taaskasutamiseks. Raamatukogu nõustab teadlasi andmehaldusplaanide väljatöötamisel, et oleks tagatud teaduse avaandmete säilimine ja taaskasutamine. Pakutakse igakülgset tuge teadustulemuste (teaduspublikatsioonid ja -andmed) avatud juurdepääsuga publitseerimisel, nõustades nii avatud sisulitsentside kui ka publitseerimisplatvormide asjus“ (Tartu Ülikooli raamatukogu..., 2015: 4)

Tartu Ülikooli raamatukogu pakub mitmeid andmehalduseks vajalikke koolitusi ja teenuseid juba aastast 2015 ning on olnud avatud juurdepääsu toetaja alates 2009. aastast (Tartu Ülikooli raamatukogu..., 2016).



## 2 TEADUSANDMETE HALDAMINE

Järgnevalt kirjeldatakse lühidalt andmehalduse põhimõtteid.

Selleks, et rahastajate ja ajakirjade poolt esitatud nõudmistele vastata ning teadusandmete jagamine ning taaskasutamine oleks võimalik, on oluline, et andmete säilimise eest oleks korralikult hoolitsetud. Seda saab tagada teadusandmete haldamisega.

Kõige üldisemalt tähendab andmete haldamine andmete eest hoolitsemist nii, et töö nendega sujub probleemivabalt. Konkreetset andmekogu ja sellega seonduvat dokumentatsiooni on kerge leida nii uurimisprojekti vältel kui ka pärast selle lõppemist. Kolleegide andmed on lihtsasti kasutatavad ja arusaadavad (Briney, 2015: 7-8).

Andmete taaskasutamiseks peavad nad sisaldama toetavat taustainformatsiooni, mis võimaldab andmeid mõista, võrrelda ja korduskatseid teha. Näiteks leiab Zimmerman oma uurimuses, et ökoloogide peamiseks kriteeriumiks andmete taaskasutamisel on võimalus andmetest aru saada, mitte see, et algandmete kogujat isiklikult tuntakse (Zimmerman, 2008: 634). Andmete mõistmiseks on vajalik varustada need metaandmetega. Oluline on märkida kuidas uurimisinstrument loodi, kus, kuidas ja millal andmeid koguti ning detailne kirjeldus andmete töötlemise protsessist (Gray, Szalay, Thakar, Stoughton, & others, 2002).

Teadusandmete haldamine hõlmab erinevaid andmetega seotud tegevusi. Näiteks andmehalduse planeerimist, andmete dokumenteerimist ja organiseerimist, analüüsimist, tundlike andmete kaitsmist, andmete majutamist ja varundamist uurimisprojekti vältel, andmete eest hoolitsemist pärast uurimisprojekti lõppu, andmete jagamist, uute projektide jaoks andmete leidmist (Briney, 2015: 7). Teadusandmete korralik haldamine võib algselt tunduda tülikas, kuid hoiab hiljem kokku teadlase aega ja närvikulu (Briney, 2015: 7). Oma andmete kogumisega saadud teadmised aitavad ka teiste kogutud andmeid paremini mõista (Zimmerman, 2008: 634).

### 2.1 Teadusandmete elutsükkel

Teadusandmete haldamise illustreerimiseks kasutatakse teadusandmete elutsükli mudelit. (Briney, 2015: 9; Corti, Van den Eynden, Bishop, & Woollard, 2014: 17-18).

Corti jt toetuvad kuuest komponendist koosnevale andmete elutsükli mudelile, mis koosneb järgmistest faasidest (Corti, Eynden, et al., 2014:17-18):

- olemasolevate andmete otsimine ja andmehalduse planeerimine (*discovery and planning*);
- andmete kogumine (*data collection*);

- andmete töötlemine ja analüüsimine (*processing and analysis*);
- andmete avaldamine ja jagamine (*publishing and sharing*);
- andmete pikaajaline säilitamine (*long-term management*);
- andmete taaskasutamine (*reusing data*).

Andmete elutsükli kõik osad on seotud erinevate toimingutega. Andmete leidmise ja planeerimise etapis toimub uurimisprojekti kavandamine, olemasolevate andmeallikate otsimine, andmehaldusplaani koostamine, otsutamine andmete jagamise võimalikkuse üle (Corti, Van den Eynden, et al., 2014: 17).

Andmete kogumise etapp on seotud erinevate meetodite abil andmete kogumisega, andmete metaandmete loomise ja hankimisega ning teiste poolt kogutud andmete hankimisega (Corti, Van den Eynden, et al., 2014: 17). Andmete töötlemise ja analüüsimise etapis tegeletakse andmete sisestamise, digiteerimise, transkribeerimise ja tõlkimisega. Samuti toimub siin andmete kontrollimine, puhastamine ja vajadusel anonümiseerimine, andmete kirjeldamine ja dokumenteerimine, andmeanalüüs ning interpreteerimine, tulemuste kirjeldamine, kasutatud andmeallikate tsiteerimine. Väga oluline on andmete järjepideva varundamise tagamine (Corti, Van den Eynden, et al., 2014: 17-18).

Andmete avaldamise ja jagamise etapis toimub andmete litsenseerimine ja autoriõiguse kehtestamine. Andmed varustatakse nende leidmist võimaldavate metaandmete ja dokumentatsiooniga ning tehakse kättesaadavaks (Corti, Van den Eynden, et al., 2014: 18). Andmete pikaajalise säilitamise etapp on seotud andmete arhiveerimisalaste tegevustega. Näiteks andmete migreerimine, varundamine ja majutamine, metaandmete ja dokumentatsiooni kogumine ja loomine ning andmete säilitamine (Corti, Van den Eynden, et al., 2014: 18). Andmete taaskasutamise etappi iseloomustavad sekundaaranalüüsid, uued uuringud, andmete kasutamine õpetamisel ja õppimisel (Corti, Van den Eynden, et al., 2014: 18).

## **2.2 Andmehaldusplaani**

Eelpool oli juttu rahastajate nõudmisest esitada granditaotlusega koos andmehaldusplaani. Andmehaldusplaani koostamine aitab läbi mõelda, kuidas andmeid hallatakse kogu uurimisprojekti vältel ja pärast seda, milline on vajadus lisaressursside järele, millised on individuaalsed ja institutsionaalsed rollid ning kohustused. Juba enne andmete kogumist tuleb välja selgitada, kas on piiranguid andmete jagamiseks, näiteks võib olla tegemist tundlike isikuandmetega (Corti, Van den Eynden, et al., 2014: 24). Kuigi hetkel on andmehaldusplaani

koostamisel peamiseks indikaatoriks rahastajate nõudmised, siis tegelikult on see, nagu kogu andmehaldus, kasulik teadlase enda jaoks.

## 2.3 Andmerepositooriumid

Pärast uurimisprojekti lõppu tuleb teadusandmed pikaajaliseks säilitamiseks, ülesleidmiseks ja juurdepääsu võimaldamiseks arhiveerida.

Andmerepositooriumid on digitaalsed arhiivid teadusandmete ja nende metaandmete kogumiseks, säilitamiseks ning tasuta kättesaadavaks tegemiseks. Sõltuvalt teadlase eelistustest või ajakirja nõudmistest, saab andmetele anda piiratud või vaba juurdepääsu (OpenAire+, 2016).

Andmerepositooriumid võivad olla institutsionaalsed, valdkonnapõhised, valdkonnaülesed (multidistsiplinaarsed) või projektipõhised (Pampel et al., 2013: 3-6).

Institutsionaalsed andmerepositooriume haldavad tavaliselt ülikoolid või teadusasutused. Kusjuures ülikooli tasemel on enamasti tegemist valdkonnaülese institutsionaalse repositooriumiga (Pampel et al., 2013: 4). Reeglina on teadusandmete arhiveerimisõigus vastava institutsiooni liikmetel. Ülikoolides haldavad repositooriume enamasti raamatukogud (OpenAire+, 2016). Institutsionaalsed andmerepositooriumid on näiteks Edinburgh DataShare (<http://datashare.is.ed.ac.uk/>) ja Tartu Ülikooli repositoorium teadusandmetele (<http://datadoi.ut.ee/>).

Valdkonnapõhised andmerepositooriumid koguvad ja säilitavad valdkonnaspetsiifilisi teadusandmeid (Pampel et al., 2013: 5). Nendes saab arhiveerida sõltumata institutsionaalsest kuuluvusest või geograafilisest asukohast (OpenAire+, 2016). Näiteks PANGAEA (<https://www.pangaea.de/>), ArXiv (<http://arxiv.org/>).

Valdkonnaülesed andmerepositooriumid arhiveerivad erinevate valdkondade teadusandmeid (Pampel et al., 2013: 6). Valdkonnaülesed andmerepositooriumid on näiteks Figshare (<https://figshare.com/>), Zenodo (<http://www.zenodo.org/>), Dryad (<http://datadryad.org/>).

Projektipõhised andmerepositooriumid on loodud kindla uurimisprojekti teadusandmete säilitamiseks (Pampel et al., 2013: 6). Näiteks Scientific Drilling Database (<http://www.scientificdrilling.org/>).

Lisaks on olemas suletud repositooriumid nende andmete jaoks, mida ei ole võimalik avalikult kättesaadavaks teha. Tavaliselt on tegemist tundlike andmetega (Whyte, 2015), näiteks UK Data Service Secure Lab (<https://www.ukdataservice.ac.uk/use-data/secure-lab/>).

### 3 UURINGU LÄBIVIIMINE

#### 3.1 Uuringu eesmärk

Töö empiirilise osa eesmärk on välja selgitada, kirjeldada ja võrrelda teadusandmete kogumise, säilitamise, kättesaadavaks tegemise ja taaskasutamise praktikad Tartu Ülikooli loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna teadlaste ja doktorantide seas.

##### **Peamised uurimisküsimused:**

- kuidas ja millist tüüpi teadusandmeid kogutakse;
- kas ja kus teadusandmeid säilitatakse;
- kuivõrd ollakse valmis oma andmeid teistega jagama ning millised on peamised takistused;
- kas ollakse teadlik andmehaldusplaani;
- kas on akadeemilisest positsioonist lähtuvaid erinevusi;
- kas on valdkondadevahelisi erinevusi;
- millisest lisainfost ollakse huvitatud.

#### 3.2 Metoodika

Uurimistöö läbiviimiseks kasutati veebipõhist ankeetküsitlust. Ankeedi koostamisel oli eeskujuks 2015. aastal Austrias läbiviidud teadusandmetealase uuringu küsimustik (Bauer, Bruno et al., 2015), mida töö autor teooriale tuginedes täiendas. Akadeemilise positsiooni määratlemise küsimuse koostamisel lähtuti dokumendist „Tartu Ülikooli akadeemilise personali ametijuhend“ (Tartu Ülikool, 2012).

Enne küsimustiku avalikustamist paluti seda testida kolmel loodus- ja täppisteaduste valdkonna doktorandil ning ühel sotsiaalteaduste valdkonna akadeemilisel töötajal ning põhjendatud juhul viidi sisse soovitatud muudatused.

Ankeet koosnes kokku 30 valikvastustega küsimusest (lisa 3). 12 küsimuse puhul võis valida mitu vastusevarianti, 9 küsimuse puhul ühe vastusevariandi ning 2 küsimust olid vaba tekstiga küsimused. Vastaja isiku kohta esitati kokku 7 küsimust. Kõigi küsimuste juures oli võimalik lisada vaba tekstina kommentaare.

Küsimustik oli jaotatud viide plokki:

- teadusandmete kogumine;
- teadusandmete arhiveerimine;
- teadusandmete kättesaadavus, nähtavus ja taaskasutatavus;

- infrastruktuur ja teenused;
- andmed vastaja kohta.

Küsimustiku täitmiseks arvestati orienteeruvalt kuni 15 minutit. Kõigile küsimustele vastamine oli vabatahtlik ning anonüümne. Küsimustik oli eestikeelne.

Andmeid koguti perioodil 9.03-29.03.2016. Esimene pöördumine palvega edastada kutse uuringus osalemiseks akadeemiliste töötajate ja doktorantide listidesse, saadeti e-posti teel loodus- ja täppisteaduste ja sotsiaalteaduste valdkonna instituutide sekretärile, õppekorraldusspetsialistidele ning Tartu Ülikooli raamatukogu erialareferentidele 9. märtsil 2016. Teine kutse saadeti välja 23. märtsil 2016.

Küsimustiku vormistamiseks, edastamiseks ja andmete kogumiseks kasutati Tartu Ülikooli UT LimeSurvey platvormi (UT - LimeSurvey, 2016). Andmeanalüüsiks kasutati programme IBM SPSS Statistics 24 ja Microsoft Excel 2010.

Tulemuste analüüsimisel kasutati kvantitatiivseid meetodeid. Küsimuste lõikes on välja on toodud kirjeldav statistika (keskmised, protsendid, risttabelid). Erinevate gruppide võrdlemisel on statistilise olulisuse kontrollimiseks kasutatud hii-ruut testi ja Crameri V testi. Välja on toodud ainult statistiliselt olulised erinevused gruppide vahel.

Vaba tekstina esitatud kommentaarid ei ole alati esitatud otsesitaatidena, osasid tsitaate on lühendatud ning vajadusel on eemaldatud vastaja isiku identifitseerimist võimaldavad viited.

### **3.3 Valim**

Kutse uuringus osalemiseks saadeti kõigile Tartu Ülikooli loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna akadeemilistele töötajatele ja doktorantidele, keda on Tartu Ülikooli personali- ja õppeosakonna andmetel 1750 (Roosimäe, 2016; Tüli, 2016) (lisa 1). Valimist jäeti välja välismaalastest akadeemilised töötajad ja doktorandid. Küsimustikule vastas kokku 391 inimest, mis teeb 22% koguvalimist.

Loodus- ja täppisteaduste valdkonda kuuluvad järgmised Tartu Ülikooli üksused (Tartu Ülikooli kodulehekülg, 2016):

- arvutiteaduse instituut;
- Eesti mereinstituut;
- füüsika instituut;
- keemia instituut;
- matemaatika ja statistika instituut;

- molekulaar- ja rakubioloogia instituut;
- tehnoloogiainstituut;
- ökoloogia ja maateaduste instituut.

Sotsiaalteaduste valdkonda kuuluvad järgmised Tartu Ülikooli üksused (Tartu Ülikooli kodulehekülg, 2016):

- haridusteaduste instituut;
- Johan Skytte poliitikauuringute instituut;
- majandusteaduskond;
- psühholoogia instituut;
- õigusteaduskond;
- ühiskonnateaduste instituut;
- Narva kolledž;
- Pärnu kolledž.

Vastanutest 234 määratles valdkonna, millega ta on seotud. Nendest 72% on seotud loodus- ja täppisteaduste valdkonnaga ning 28% sotsiaalteaduste valdkonnaga.

40% kõigist vastanutest jätsid oma valdkonna määratlemata. Arvestades, et küsimustik oli adresseeritud ainult loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna akadeemilistele töötajatele ja doktorantidele ning eeldades, et tegemist on just nende valdkondadega seotud inimestega, kelle arvamus on uuringu teema osas oluline, on analüüsitud ka nende isikute vastuseid, kes valdkonda ei märkinud.

Valdkondli võrdlevalt analüüsid on arvestatud ainult nende vastajatega, kes valdkonna määratlesid.

Vastanutest 227 täpsustas instituudi, millega ta on seotud. Loodus- ja täppisteaduste valdkonnas oli kõige rohkem vastajaid ökoloogia ja maateaduse instituudist ning sotsiaalteaduste valdkonnas ühiskonnateaduste instituudist (lisa 2).

Vastanutest 227 määratles oma akadeemilise positsiooni, nendest 61% on akadeemilised töötajad ning 39% on doktorandid. 42% kõigist vastanutest jättis akadeemilise positsiooni täpsustamata.

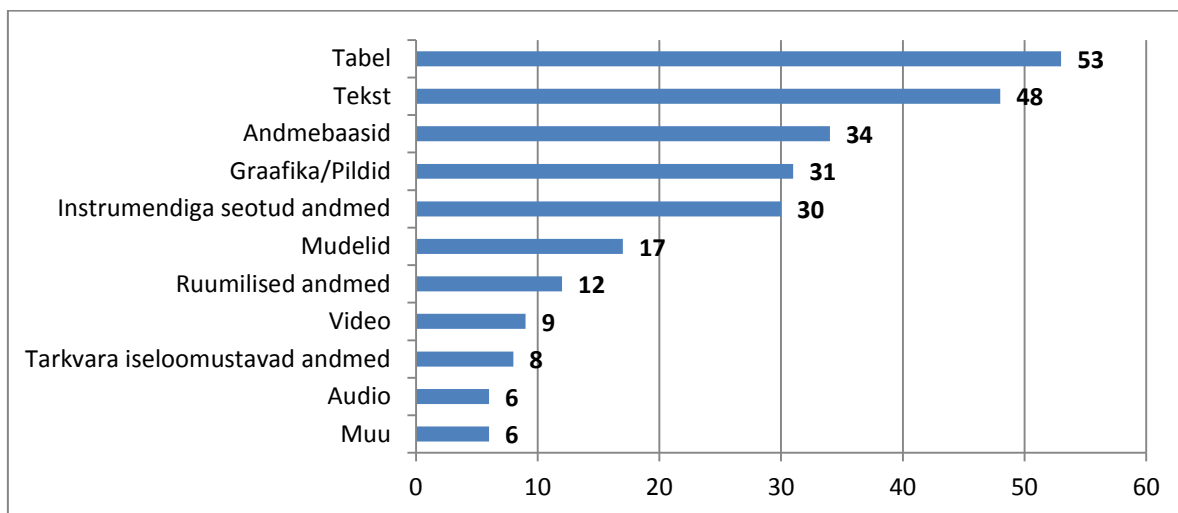
Vastajate keskmine vanus on 38 aastat ( standardhälve 12 aastat).

## 4 TULEMUSTE ANALÜÜS

### 4.1 Teadusandmete kogumine

Esimese küsimuste ploki eesmärk oli välja selgitada, kuidas ning milliseid teadusandmeid Tartu Ülikooli loodus- ja täppis- ning sotsiaalteadlaste ja doktorantide poolt kogutakse.

Küsimusele, millist tüüpi on algandmed, mida uurimistöö käigus kogute, võis valida mitu vastusevarianti. Kõige rohkem esinevad uurimistöö käigus kogutud andmed tabeli (53%) ja teksti (48%) kujul (joonis 1). Järgnevad andmebaasidena (34%) ja graafiliselt või pildina (31%) kogutud andmed.



**Joonis 1. Uurimistöö käigus kogutavate algandmete tüüp (%)**

Lisakommentaaries täpsustati, et kogutakse *teksti kujul esitatavad DNA ja RNA sekventeerimisandmed ning numbrilisi mõõteandmed*.

Loodus- ja täppisteaduste valdkonnas on esimesel kohal tabeli kujul kogutavad andmed (tabel 1). Väga harva on andmed esitatud audiona. Üle poolte kogutavatest andmetest on lisaks tabelitele, tekstid, andmebaasid, graafika/pildid ning instrumentidega seotud andmed.

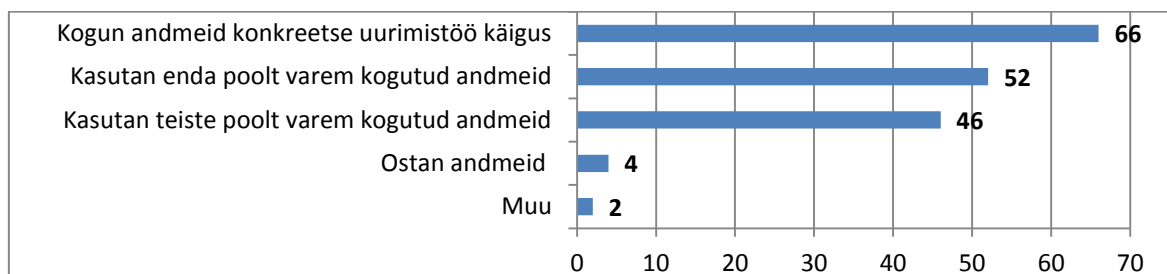
Sotsiaalteaduste valdkonnas kogutakse kõige enam andmeid teksti kujul (tabel 1) ning üle poole andmetest kogutakse tabelina.

**Tabel 1. Uurimistöö käigus kogutavate algandmete tüüp. Valdkondlik jaotus**

Kogutavad andmed	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Tekst	86	63
Tabel	57	85
Audio	28	1
Video	19	12
Graafika/Pildid	22	53
Andmebaasid	37	57
Instrumendiga seotud andmed	23	51
Mudelid	12	28
Ruumilised andmed	3	24
Tarkvara iseloomustavad andmed	5	15

\*\*\* $p < 0,000$

Küsimusele, kuidas uurimistöö jaoks andmeid saadakse, võis valida mitu vastusevarianti. Konkreetse uurimistöö jaoks kogub andmeid selle sama uurimistöö käigus 66% vastanutest (joonis 2). Enda poolt varem kogutud andmeid kasutab 52% vastanutest. Teiste poolt kogutud tasuta kättesaadavaid andmeid kasutab 46% vastanutest. Lisakommentaaries toodi välja, et *kasutakse uurimisgrupi liikmete poolt varem kogutud andmeid, mis ei ole tasuta kättesaadavad ega avalikud ning aparatuur kogub andmeid automaatselt.*

**Joonis 2. Kuidas saate uurimistöö jaoks andmeid? (%)**

Nii loodus- ja täppisteaduste kui ka sotsiaalteaduste valdkonnas koguvad pea kõik vastanud oma uurimistöö jaoks andmeid konkreetse uurimistöö käigus (tabel 2). Kolm neljandikku vastanutest kasutab ka enda poolt varem kogutud andmeid. Loodus- ja täppisteaduste valdkonnas kasutatakse pisut rohkem (70%) teiste poolt kogutud tasuta kättesaadavaid andmeid kui sotsiaalteaduste valdkonna vastanute hulgas (62%)

**Tabel 2. Kuidas saate uurimistöö jaoks andmeid? Valdkondlik jaotus.**

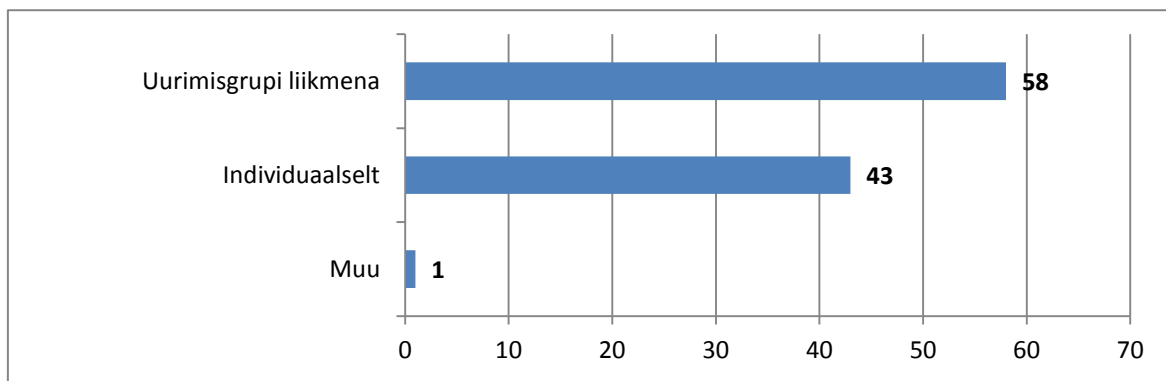
Kuidas saate uurimistöö jaoks andmeid	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Kogun andmeid konkreetse uurimistöö käigus	99	95
Kasutan enda poolt varem kogutud andmeid	74	77
Kasutan teiste poolt varem kogutud andmeid	62	70

\*\*\* $p < 0,000$



Vastuseks küsimusele, kas kogute teadusandmeid individuaalselt või uurimisgrupi liikmena, võis valida mitu vastusevarianti. 58% vastanutest kogub andmeid uurimisgrupi liikmena ning 43% individuaalselt (joonis 3). Mitmel korral täpsustati et, *kasutatakse sama uurimisgrupi poolt varem kogutud avalikustamata andmeid*.

**Joonis 3. Kas andmeid kogutakse individuaalselt või uurimisgrupi liikmena?**



Valdkondadevahelise erinevusena on loodus- ja täppisteaduste puhul täheldatav suurem osakaal uurimisgrupis andmete kogumisel (tabel 3). Sotsiaalteaduste valdkonnas kogutakse andmeid peaaegu võrdselt nii individuaalselt (77%) kui ka uurimisgrupis (74%).

**Tabel 3. Kas andmeid kogutakse individuaalselt või uurimisgrupi liikmena?**

**Valdkondlik jaotus**

Kas kogute andmeid individuaalselt või uurimisgrupis	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Individuaalselt	77	57
Uurimisgrupi liikmena	74	88

\*\*\* $p < 0,000$

Vastuseks küsimusele, mitme uurimisprojektiga olite seotud eelmisel aastal, sai valida valikvastustest ühe. Sellele küsimusele vastas 70% kõigist uuringus osalenutest. Enamus vastanutest oli seotud ühe uurimisprojektiga (tabel 4). 2-5 projektiga oli seotud 10% vastanutest. 22% vastajatest valisid vastuseks muu.

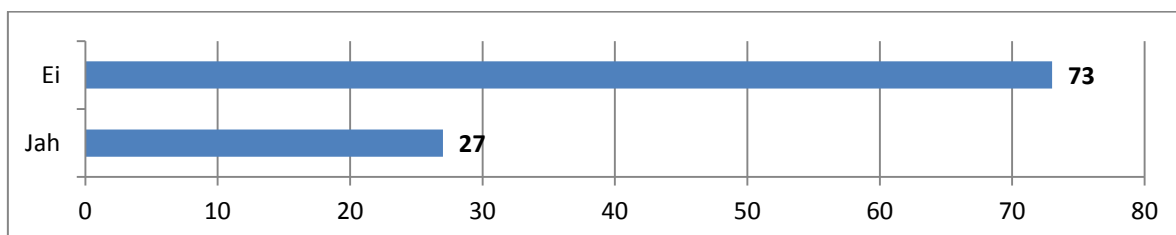
**Tabel 4. Mitme uurimisprojektiga olite seotud eelmisel aastal? (%)**

Mitme uurimisprojektiga olite seotud eelmisel aastal	%
5 ja rohkem	1
0	6
2-5	10
Muu	22
1	61

## 4.2 Teadusandmete arhiveerimine

Teise küsimuste ploki eesmärk oli välja selgitada teadusandmete säilitamise ja arhiveerimisalsed praktikad uurimisprojekti vältel ja pärast selle lõppemist.

Küsimusele, kas teil on teadusandmeid kaduma läinud, sai valida ühe vastusevariandi. Sellele küsimusele vastas 67% kõigist vastanutest. Üle veerandi vastajatest märkis, et neil on teadusandmeid kaduma läinud (joonis 4). Lisakommentaaries on andmekao põhjustena välja toodud: *arvuti või mälupulga rike; enda või teiste hooletus; oskamatus andmeid tähistada nii, et need oleks hiljem leitavad; mälupulga vargus; paberkandjal ankeetide ja laborimärkmete kadumine.*



**Joonis 4. Kas teil on teadusandmeid kaduma läinud? (%)**

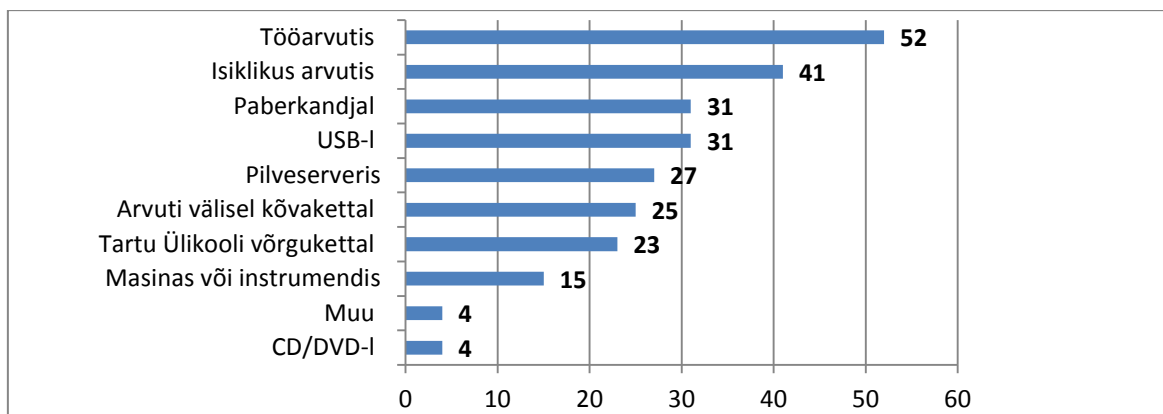
Andmekaoga on kokku puutunud 31% loodus- ja täppisteaduste valdkonna vastanutest ja 17% sotsiaalteaduste valdkonna vastanutest (tabel 5).

**Tabel 5. Kas teil on teadusandmeid kaduma läinud? Valdkondlik jaotus**

Kas teil on teadusandmeid kaduma läinud?	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Jah	17	31
Ei	83	69

\*\*\* $p < 0,000$

Vastuseks küsimusele, kus hoiate tavaliselt oma teadusandmeid uurimisprojekti vältel, võis valida mitu vastusevarianti. 52% vastanutest hoiab oma andmeid uurimisprojekti ajal tööarvutis ning 41% vastanutest isiklikus arvutis (joonis 5). USB on valinud andmete hoidmiseks 31% vastanutest. Pilveserveris hoiab andmeid 27%, arvuti välisel kõvakettal 25% ning Tartu Ülikooli võrgukettal 23% vastanutest. Ühtlasi selgus, et 31% teadusandmetest hoitakse uurimisprojekti vältel paberkandjal. Lisakommentaaries toodi välja andmete säilitamine *instituudi serveris; välises arvutiklastris; Google'i pakutavas keskkonnas; diktofonil; paberitel juhendaja lukustatud kapis.*



**Joonis 5. Kus hoitakse teadusandmeid uurimisprojekti vältel? (%)**

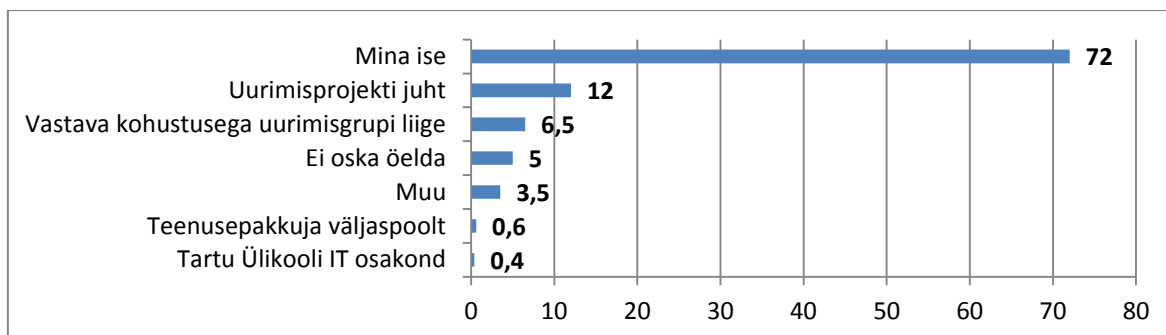
Valdkondadevahelise erinevusena saab välja tuua, et loodus- ja täppisteaduste vastanud hoiavad andmeid pigem tööarvutis kui isiklikus arvutis. Sotsiaalteaduste valdkonna vastanud kasutavad isiklikku ja tööarvutit andmete hoidmiseks peaaegu võrdselt (tabel 6). Märkimisväärne on, et 51% loodus- ja täppisteaduste andmetest on paberkandjal. Pea pooled vastanud nii loodus- ja täppisteaduste kui ka sotsiaalteaduste valdkonnast kasutavad andmete hoidmiseks USB-mälupulka.

**Tabel 6. Teadusandmete hoidmine uurimisprojekti vältel. Valdkondlik jaotus**

Andmete hoidmine uurimisprojekti vältel	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Isiklikus arvutis	66	60
Tööarvutis	62	88
Pilveserveris	37	41
Tartu Ülikooli võrgukettal	19	37
Arvuti välisel kõvakettal	32	41
USB-l	45	47
CD/DVD-l	5	6
Masinas või instrumendis	5	31
Paberkandjal	37	51
Muu	8	6

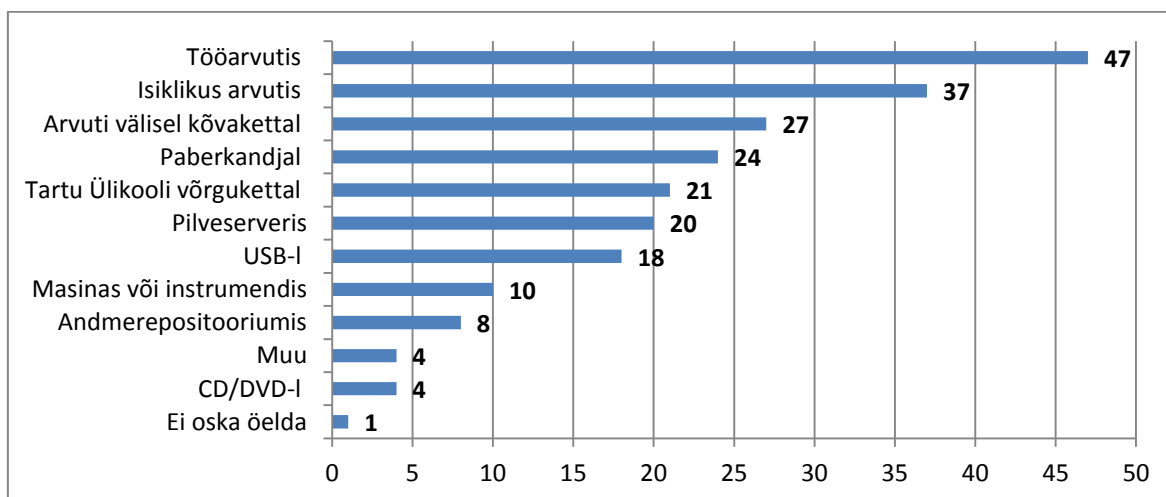
\*\*\* $p < 0,000$

Vastuseks küsimusele, kes vastutab enamasti uurimisprojekti vältel teadusandmete säilitamise eest, sai valida ühe vastusevariandi. Ligi kolmveerand vastanutest näeb vastutavana iseennast (joonis 6). Umbes veerand vastanutest arvab, et uurimisprojekti vältel vastutab andmete säilimise eest uurimisprojekti juht.



**Joonis 6. Uurimisprojekti vältel teadusandmete säilitamise eest vastutaja (% , n=261)**

Vastuseks küsimusele, kus hoiate tavaliselt oma teadusandmeid pärast uurimisprojekti lõppu, võis valida mitu vastusevarianti. Tööarvutis hoiab pärast uurimisprojekti lõppu andmeid 47% ning isiklikus arvutis 37% vastanutest (joonis 7). Arvuti välist kõvaketast kasutab uurimisprojekti järgselt andmete hoidmiseks 27%, Tartu Ülikooli välist võrguketast 21% ning pilveserverit 20% vastanutest. Lisakommentaaris märgiti: *välises arvutiklastris; instituudi serveris; andmed hävitatakse pärast uurimisprojekti lõppu*.



**Joonis 7. Kus hoitakse teadusandmeid pärast uurimisprojekti lõppu? (%)**

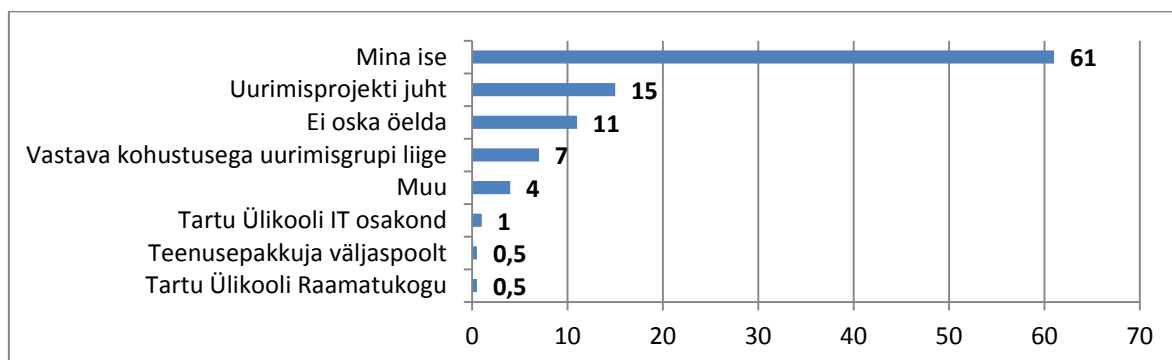
Sarnaselt uurimisprojekti vältel andmete säilitamisele eelistab 82% (tabel 7) loodus- ja täppisteaduste valdkonna vastanutest tööarvutit ning 54% sotsiaalteaduste valdkonna vastanutest isiklikku arvutit uurimisandmete säilitamiseks ka pärast uurimisprojekti lõppu.

**Tabel 7. Teadusandmete hoidmine pärast uurimisprojekti lõppu. Valdcondlik jaotus**

Andmete hoidmine pärast uurimisprojekti lõppu	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Andmerepositooriumis	9	14
Isiklikus arvutis	54	55
Töõarvutis	49	82
Pilveserveris	31	30
Tartu Ülikooli võrgukettal	20	33
Arvuti välisel kõvakettal	35	43
USB-l	32	26
Masinas või instrumendis	3	20
Paberkandjal	26	40

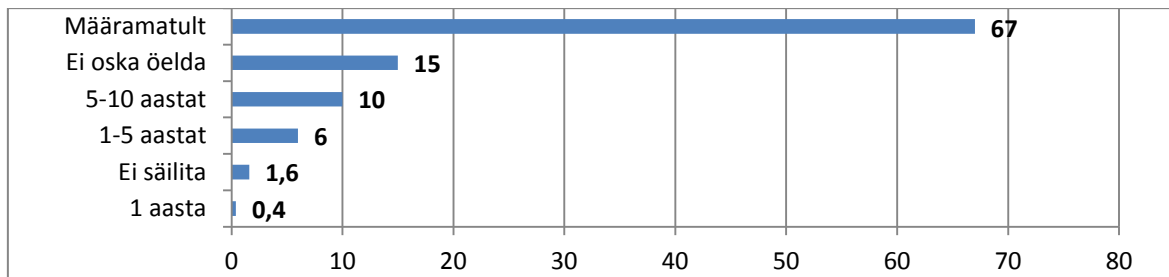
\*\*\* $p < 0,000$

Vastuseks küsimusele, kes vastutab enamasti pärast uurimisprojekti lõppu teadusandmete arhiveerimise eest, võis valida ühe vastusevariantidest. 61% küsimusele vastanutest arvab, et nemad ise vastutavad arhiveerimise eest (joonis 8). 15% peab vastutavaks uurimisprojekti juhti ja 11% ei oska öelda. Lisakommentaaries täpsustati: *ei arhiveeri, säilitan vaid artiklite väljatrükid; mitte keegi ei vastuta; praegune süsteem toimib ühe inimese idealistlikul visioonil*.



**Joonis 8. Pärast uurimisprojekti lõppu teadusandmete säilitamise eest vastutaja (% , n=262)**

Vastuseks küsimusele, kui kaua plaanite säilitada teadusandmeid pärast uurimisprojekti lõppu, sai valida ühe vastusevariandi. Määramatult säilitab teadusandmeid 67% vastanutest (joonis 9). 15% vastanutest ei oska öelda kaua peaks andmeid säilitama.

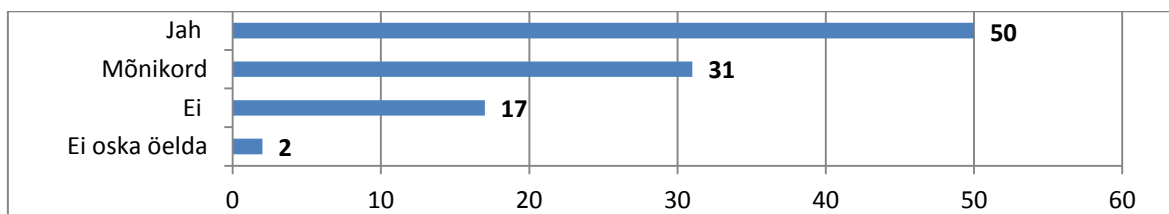


**Joonis 9. Uurimisprojektijärgne teadusandmete säilitamise periood (% , n=262)**

### 4.3 Teadusandmete kättesaadavus, nähtavus ja taaskasutatavus

Kolmanda küsimuste ploki eesmärk oli välja selgitada kas ja kuidas ollakse valmis oma teadusandmeid jagama.

Pooled vastanutest arvasid, et nende andmed võiksid tulevikus teistele huvi pakkuda (joonis 10).



**Joonis 10. Kas kogutud teadusandmed võiksid tulevikus teistele huvi pakkuda? (% , n=248)**

Lisakommentaaries toodi välja, et et andmete jagamine võimaldab teadmiste akumulatsiooni, mis pikemas perspektiivis viib täpsema teadmiseni: („Mida rohkem andmeid erinevatest maailmajagudest, seda täpsemad algoritmid saab välja töötada“, „Andmed võivad olla äärmiselt huvi pakkuvad teadlastele, kes uurivad samasid objekte, mida mina olen mõõtnud ja analüüsinud. Pikemas perspektiivis on väga paljud astrofüüsikalised objektid huvipakkuva(ma)d, kui nende kohta on olemas pikad aegread“, „DNA järjestuseandmed on seeneökoloogias järjest laialdasemalt kasutusel teiste sama valdkonna uurijate poolt“) ning loodeti, et andmete jagamisest tõuseb üleüldist kasu („Ma ei teed ju tööd töö tegemise pärast. Ikka loodan, et kunagi oleks kellelgi sellest kasu“. Toodi välja ka teatavat skepsist kellegi teise poolt kogutud andmete suhtes: („Andmeid genereeritakse rohkem, kui keegi eales läbi suudab vaadata. Kolmanda poole eksperimentaalsete andmete suhtes on ikka teatav skepsis ning alati pole tulemused üle kantavad“). Samuti leiti, et andmete jagamist võib segada see, et andmed vananevad liiga kiiresti ning, et andmed on konfidentsiaalsed.

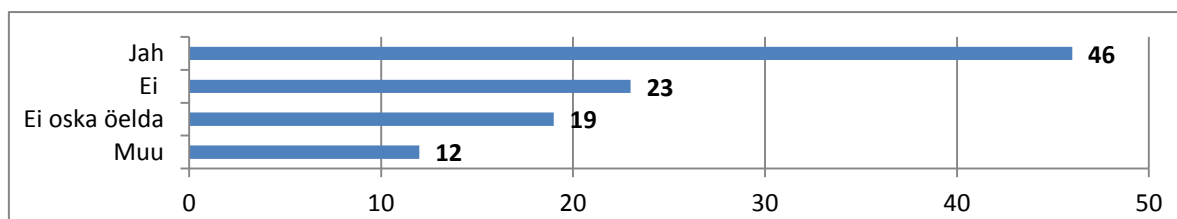
Üle poole loodus- ja täppisteaduste valdkonna vastanutest olid kindlad, et nende kogutud andmed võiksid ka tulevikus teistele huvi pakkuda (tabel 8). Sotsiaalteaduste valdkonnas arvas sama 38% vastanutest.

**Tabel 8. Kas kogutud teadusandmed võiksid tulevikus teistele huvi pakkuda? Valdkondlik jaotus**

Kas kogutud teadusandmeid pakuksid teistele huvi?	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Jah	38	56
Ei	3	1
Mõnikord	39	29
Ei oska öelda	20	14

\*\*\* $p < 0,000$

Vastuseks küsimusele, kas teadusandmed on varustatud metaandmete ja kirjeldustega nii, et need on hiljem arusaadavad ja kasutatavad teile endale ja vajadusel ka kõigile teistele väljaspool uurimisgruppi, võis valida ühe vastusevariandi. Metaandmetega varustab oma teadusandmed 46% vastanutest (joonis 11). Metaandmeid ei lisa andmetele 23% vastanutest.



**Joonis 11. Teadusandmete varustamine metaandmetega (% , n=247)**

Kommentaaries lisati, et *mõnikord on, mõnikord ei ole ning enamasti on, aga aru saan neist ise või uurimisgrupi liikmed* ning, et *kasutatakse oma süsteemi*.

Sotsiaalteaduste valdkonnas lisab metaandmeid 39% vastanutest ning loodus- ja täppisteaduste valdkonnas 48% vastanutest (tabel 9).

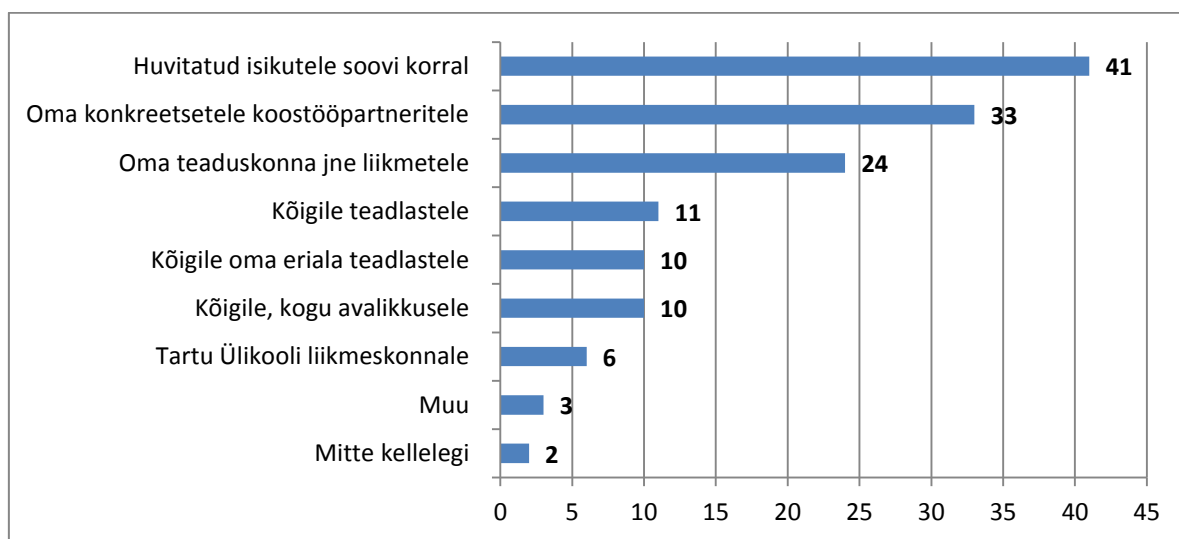
**Tabel 9. Teadusandmete varustamine metaandmetega. Valdkondlik jaotus**

Kas teadusandmed on varustatud metaandmetega	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Jah	39	48
Ei	32	20
Ei oska öelda	20	18
Muu	9	14

\*\*\* $p < 0,000$

Vastuseks küsimusele, kellele võimaldate juurdepääsu oma teadusandmetele, kui nende jagamine ei ole seotud piirangutega, võis valida mitu vastusevarianti. 41% vastanutest

annaks juurdepääsu oma andmetele juhul, kui selleks soovi avaldatakse. 33% vastanutest annaks juurdepääsu konkreetsetele koostööpartneritele ning 24% oma teaduskonna/instituudi/ uurimiskeskuse liikmetele (joonis 12).



**Joonis 12. Kellele võimaldatakse juurdepääs oma teadusandmetele? (%)**

Valdkondadevahelise suurima erinevusena ilmnes, et loodus- ja täppisteaduste valdkonna vastanud on altimad jagama oma andmeid kogu avalikkusega (tabel 10). Sotsiaalteaduste valdkonna vastanute hulgas on selleks valmis 5% vastanutest.

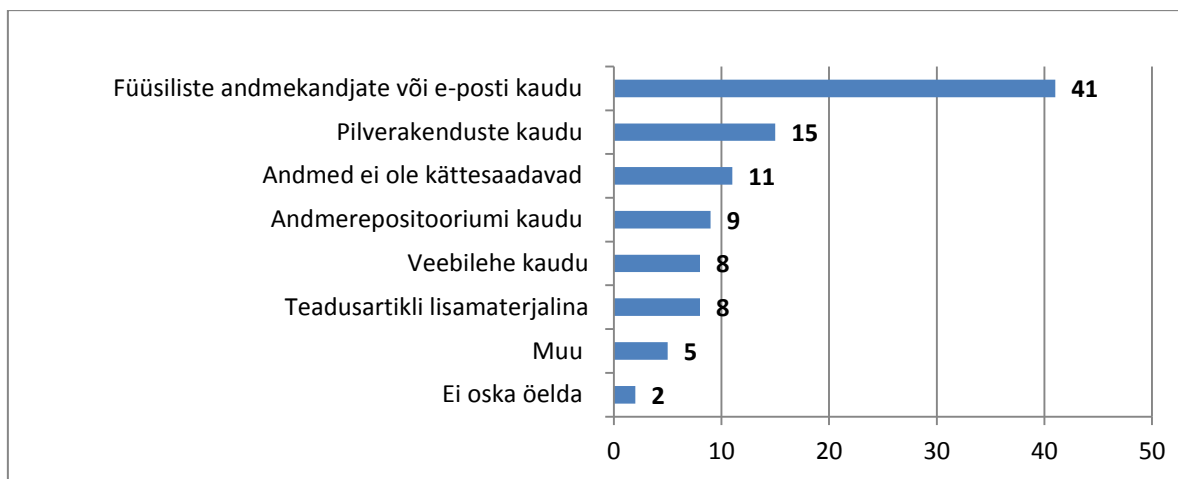
**Tabel 10. Kellele võimaldatakse juurdepääs oma teadusandmetele. Valdkondlik jaotus**

Kellele võimaldatakse juurdepääs oma andmetele?	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Kõigile, kogu avalikkusele	5	20
Kõigile teadlastele	15	16
Kõigile oma eriala teadlastele	15	17
Tartu Ülikooli liikmeskonnale	8	9
Oma teaduskonna/instituudi/uurimiskeskuse liikmetele	29	41
Oma konkreetsetele koostööpartneritele	48	52
Huvitatud isikutele soovi korral	57	67
Mitte kellelegi	8	1

\*\*\* $p < 0,000$

Vastuseks küsimusele, kuidas saavad teised kasutada teie teadusandmeid, võis valida mitu vastusevarianti. 41% vastanu teadusandmeid saab jagada füüsiliste andmekandjate või e-posti kaudu (joonis 13). Pilverakenduste kaudu saab kasutada 15% vastanu teadusandmeid. 11% andmed ei ole teistele kättesaadavad.





**Joonis 13. Kuidas saavad teised kasutada teie teadusandmeid? (%)**

Lisakommentaaries täpsustati, et andmeid jagatakse isiklikult mälupulgalt arvutisse. Mainiti ka Skype teel andmete jagamist. Samuti kasutatakse uurimisgrupis välja töötatud lahendusi: („Töörühma andmebaasilahenduste kaudu“) ja, et („Uuriija saab parooli, millega pääseb talle lubatud andmetele ligi“).

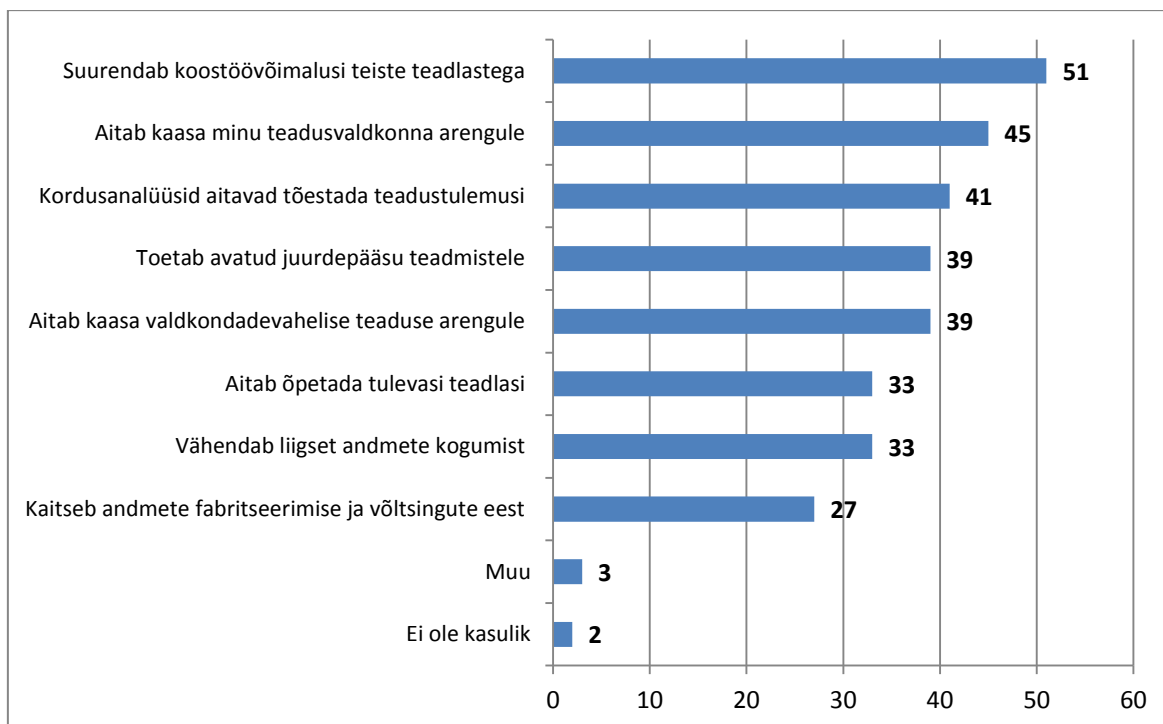
Suurima valdkondadevahelise erinevusena võib välja tuua asjaolu, et lingituna repositooriumist lisamaterjalina teadusartikli juures on andmed kättesaadavad 17% loodus- ja täppisteaduste valdkonna ja 3% sotsiaalteaduste valdkonna vastanu puhul (tabel 11).

**Tabel 11. Kuidas saavad teised kasutada teie teadusandmeid? Valdkondlik jaotus**

Kuidas saavad teised kasutada teie andmeid?	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Andmerepositooriumi kaudu	12	16
Teadusartikli lisamaterjalina	3	17
Veebilehe kaudu	8	14
Pilverakenduste kaudu	17	26
Füüsiliste andmekandjate või e-posti kaudu	57	69
Andmed ei ole kättesaadavad	28	14

\*\*\* $p < 0,000$

Vastuseks küsimusele, miks on teaduse avaandmete jagamine teie arvates kasulik, võis valida mitu vastusevarianti. 51% vastanutest leiab, et avaandmete jagamine suurendab koostöövõimalusi teiste teadlastega (joonis 14). 45% vastanu arvates aitab avaandmete jagamine kaasa tema teadusvaldkonna arengule. 41% vastanutest leiab, et kordusanalüüsid aitavad tõestada teadustulemusi.



**Joonis 14. Teaduse avaandmete jagamise kasulikkus (%)**

Lisakommentaaries toodi välja võimalus uuteks avastusteks: („*Avaldatud andmetest võib olla võimalik teha uusi tähelepanekuid ja avastusi, ilma aja ja ressursikulukat eksperimenti läbi viimata, mida nende algsed kogujad üldse ei märganud*“) ning juhtiti tähelepanu, et: („*Toetab loodushoidu ja keskkonna-alast teadlikkust*“). Samuti oldi skeptiline: („*Ei ole huvilisi, kes soovivad teiste andmetes ja mustandites kaevata*“).

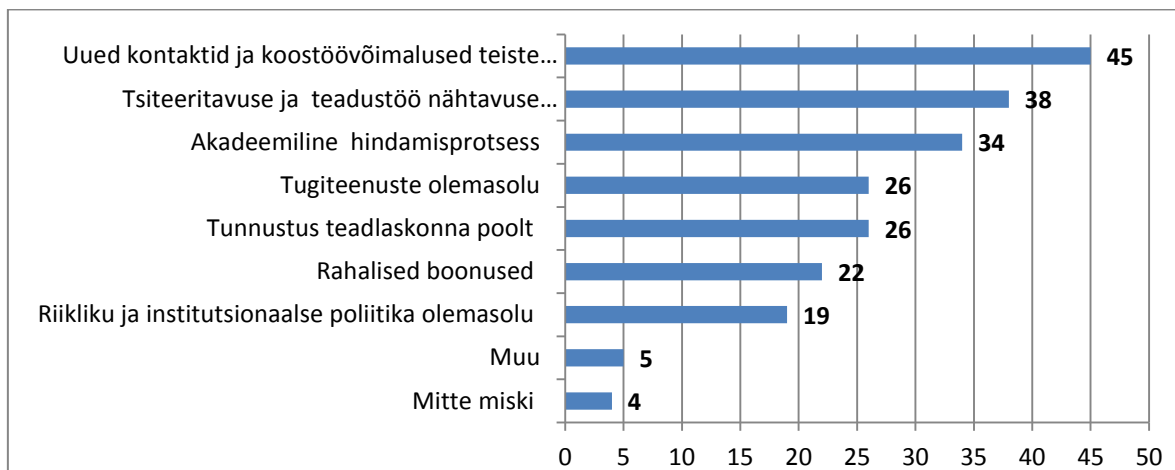
Sotsiaalteaduste valdkonna vastanutest 62% vastas, et avaandmete jagamine vähendab liigset andmete kogumist (tabel 12). Loodus- ja täppisteaduste valdkonna puhul arvas sama 47% vastanutest. Ligi kolmveerand loodus- ja täppisteaduste valdkonna vastajatest arvas, et avaandmete jagamine kaitseb fabritseerimise ja võltsingute eest. Sotsiaalteaduste valdkonnas leidis sama 40% vastanutest.

**Tabel 12. Teaduse avaandmete jagamise kasulikkus. Valdkondlik jaotus**

Miks on avaandmete jagamine kasulik	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Aitab kaasa minu teadusvaldkonna arengule	69	72
Aitab kaasa valdkondadevahelise teaduse arengule	63	61
Suurendab koostöövõimalusi teiste teadlastega	83	81
Vähendab liigset andmete kogumist	62	47
Toetab avatud juurdepääsu teadmistele	60	61
Kordusanalüüsid aitavad tõestada teadustulemusi	63	65
Kaitseb andmete fabritseerimise ja võltsingute eest	40	72
Aitab õpetada tulevasi teadlasi	54	53

\*\*\* $p < 0,000$

Vastuseks küsimusele, mis motiveeriks jagama oma teadusandmeid avatud teaduse põhimõttel, võis valida mitu vastusevarianti. 45% vastanu puhul oleks motivaatoriks uued kontaktid ja koostöövõimalused teiste teadlastega (joonis 15). 38% vastanutest motiveeriks tsiteeritavuse suurenemine ning teadustöö nähtavuse ja mõjufaktori kasv. 34% vastanutest jagaks andmeid, kui akadeemilises hindamisprotsessis arvestataks teadusandmetega kui olulise teadusliku väljundiga.



**Joonis 15. Motivaatorid teadusandmete jagamiseks avatud teaduse põhimõttel (%)**

Lisakommentaaries toodi välja: „Garantii, et teised minu andmetega uurimust kiiremini ja ilma minuta ei publitseeri“ ning „Teadusandmed on oma olemuselt avalikud andmed hetkest, mil nende tegijad on oma analüüsid teinud ja andmed avaldanud“, „Teadus ongi avatud, ei leia, et oleks vaja eraldi motivaatorit“. Samuti toodi välja murekohad: „Tugiteenused võiksid küll olla mõnikord paremad, suuremahuliste andmete avalikuks tegemine tekitab päris palju (tehnilistest raskustest põhjustatud) stressi ja ajakulu, sõltumata sellest, et tegemist on rahvusvaheliste tippasutustega, kes andmepositoriume haldavad“, „Andmete jagamine ei ole imevits“.

Üle poole loodus- ja täppisteaduste valdkonna vastanuid motiveeriks teadusandmeid jagama, kui sellega arvestataks ka akadeemilises hindamisprotsessis (tabel 13). Sotsiaalteaduste valdkonna vastanuid motiveeriks see 43% vastanute puhul.

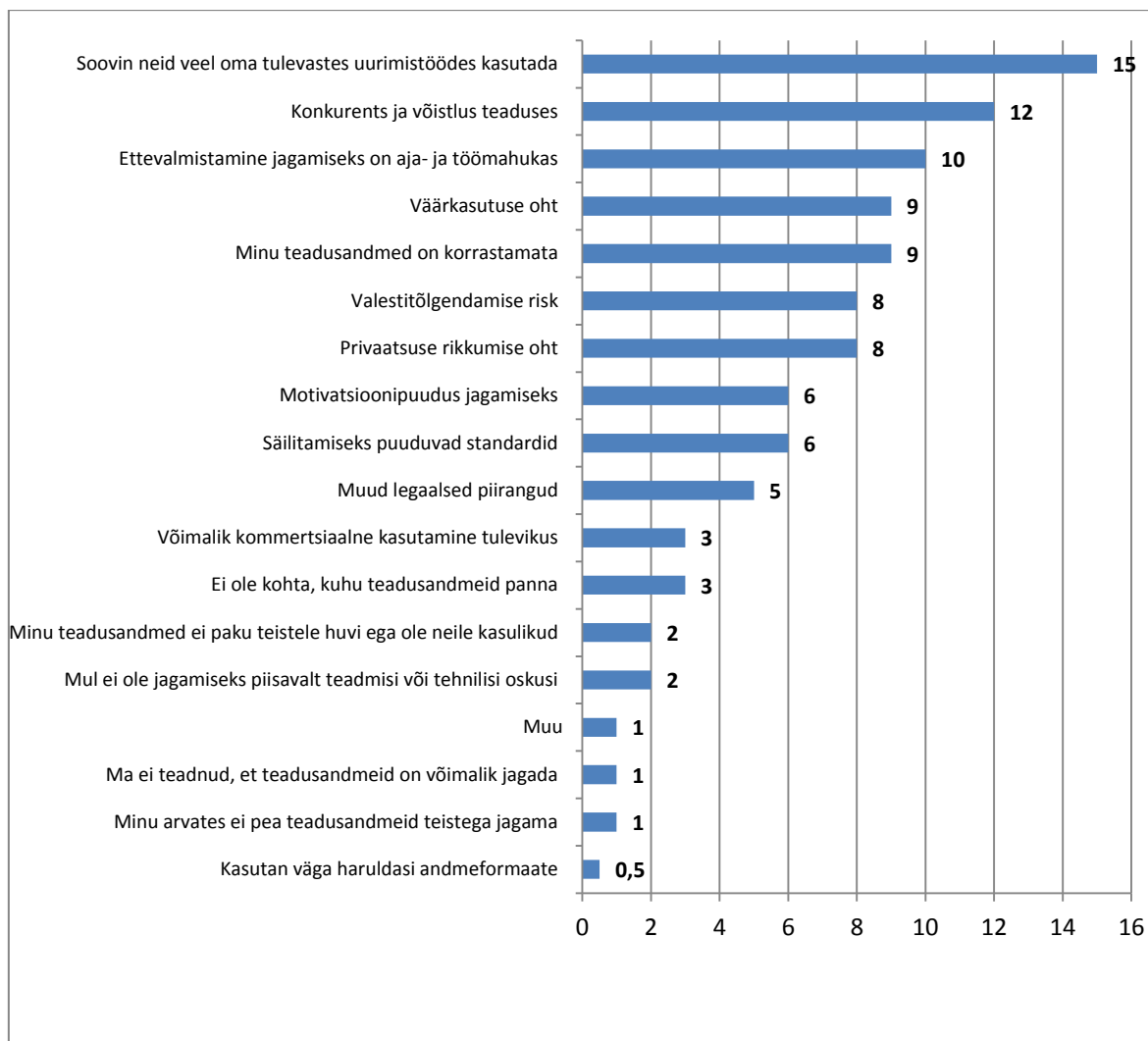
**Tabel 13. Motivaatorid teadusandmete jagamiseks avatud teaduse põhimõttel. Valdcondlik jaotus**

<b>Motivaatorid teadusandmete jagamiseks</b>	<b>Sotsiaalteadused (%)</b>	<b>Loodus- ja täppisteadused (%)</b>
Tunnustus teadlaskonna poolt	37	42
Akadeemiline hindamisprotsess	43	58
Tsiteeritavuse ja teadustöö nähtavuse suurenemine	55	63
Uued kontaktid ja koostöövõimalused teiste teadlastega	63	72
Rahalised boonused	34	34
Tugiteenuste olemasolu	45	40
Riikliku ja institutsionaalse poliitika olemasolu	32	28

\*\*\* $p < 0,000$

Kontrollküsimusele, kas jagate meeleldi oma teadusandmeid, sai valida jah või ei vastuse. Küsimusele vastas 63% kõigist vastanutest. 68% vastanutest jagab meeleldi oma teadusandmeid. 32% vastanutest ei jaga oma teadusandmeid hea meelega ning nad suunati edasi täpsustava küsimuse juurde. Valdcondlikult jagab meeleldi andmeid 69% loodus- ja täppisteaduste valdkonna vastanutest ja 57% sotsiaalteaduste valdkonna vastanutest.

Vastuseks küsimusele, mis takistab teid teadusandmeid jagamast, võis valida mitu vastusevarianti. Peamisteks andmete mittejagamise põhjuseks oli soov neid oma tulevastes uurimistöodes kasutada ja konkurents teaduses (joonis 16). Kommentaarides toodi välja, et *tegeletakse tundlike andmetega, mille avalikult jagamine ei ole lubatud.*



**Joonis 16. Teadusandmete jagamise takistused (%)**

Suurima valdkondadevahelise erinevusena ilmnnes, et sotsiaalteaduste valdkonna 28% vastanu jaoks on takistuseks privaatsuse rikkumise oht (tabel 14). Loodus- ja täppisteaduste valdkonna puhul takistaks see 9% vastanut oma andmeid jagamast. Oht, et andmeid tõlgendatakse valesti takistab 22% sotsiaalteaduste valdkonna vastanut andmeid jagamast. Loodus- ja täppisteaduste puhul takistab see 11% vastanut.

**Tabel 14. Teadusandmete jagamise takistused. Valdkondlik jaotus**

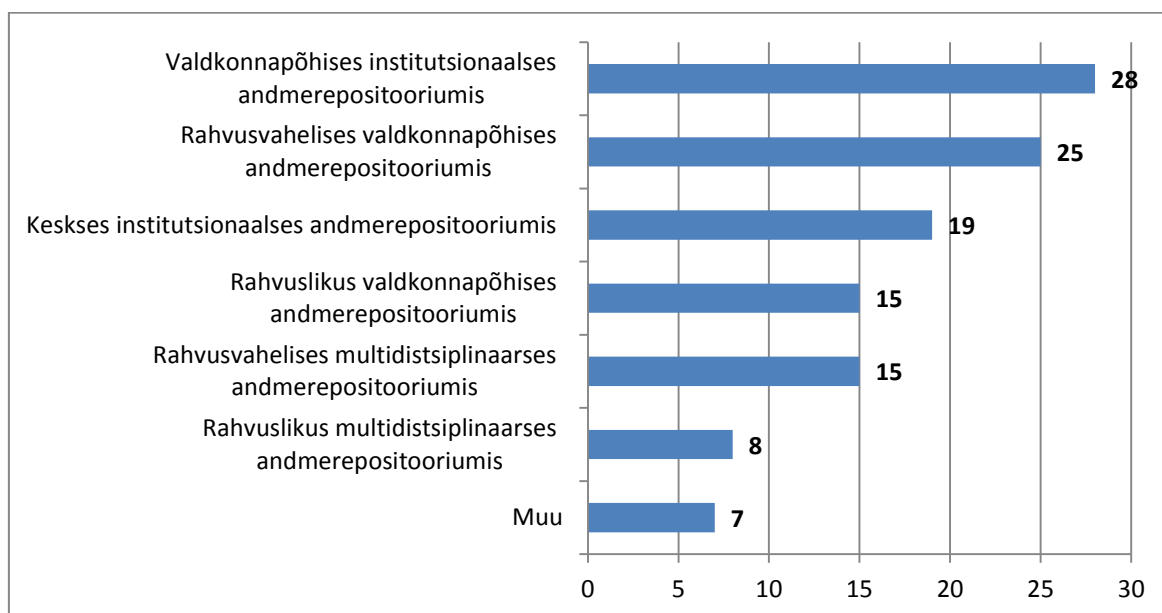
Teadusandmete jagamise takistused	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Minu teadusandmed on korrastamata	19	13
Soovin neid veel oma tulevastes uurimistöödes kasutada	28	23
Minu teadusandmed ei paku teistele huvi	10	2
Privaatsuse rikkumise oht	28	9
Teadusandmete väärkasutuse oht	19	14
Valestitõlgendamise risk	22	11
Konkurents ja võistlus teaduses	15	21
Ettevalmistamine jagamiseks on aja- ja töömahukas	20	16
Teadusandmete säilitamiseks puuduvad standardid	12	10

\*\*\* $p < 0,000$

#### 4.4 Infrastruktuur ja teenused

Neljanda küsimuste ploki eesmärgiks oli uurida vastajate teadlikkust andmerepositooriumidest ja andmehaldusplaanist ning välja selgitada, millist teadusandmetega seonduvat lisainfot vajatakse.

Vastuseks küsimusele, kus arhiveeriksid kõige meelsamini oma teadusandmeid, võis valida mitu vastusevarianti. Valdkonnapõhise institutsionaalse andmerepositooriumi valiks arhiveerimiseks 28% vastanutest (joonis 17). 25% vastanutest arhiveeriks rahvusvahelises valdkonnapõhises andmerepositooriumis.



**Joonis 17. Kus arhiveeriksid oma teadusandmeid? (%)**

Liskommentaarides täpsustati: „Ei archiveeriks üheski neist“, „DropBox, sest seal saan ise formaadi defineerida ja ei pea täitma totraid välju“, „Probleemiks (sh lisakuluks) on relatsioonilistes baasides olevate andmete teisendamine praegustesse repositooriumitesse sobivasse vormi“, „Repositooriumis, kus on võimalik erinevatele andmetele seada erinevad piirangud“ ning „Serveris ja isiklikul SSD kettal“.

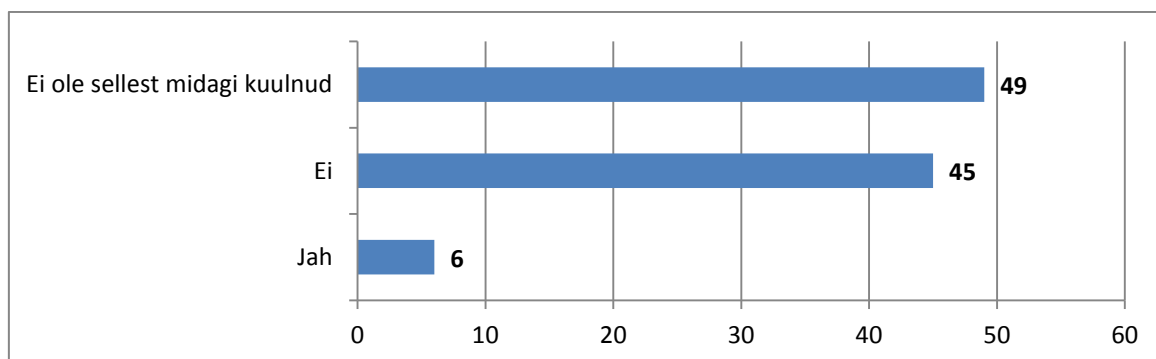
Rahvuslikus valdkonnapõhises andmerepositooriumis archiveeriks oma teadusandmeid 35% sotsiaalteaduste valdkonna vastanutest ning 20% loodus- ja täppisteaduste valdkonna vastanutest (tabel 15). Seevastu rahvusvahelises valdkonnapõhises andmerepositooriumis 46% loodus- ja täppisteaduste valdkonna vastanutest ja 31% sotsiaalteaduste valdkonna vastanutest.

**Tabel 15. Kus archiveeriksid oma teadusandmeid. Valdkondlik jaotus**

Kus archiveeriksid teadusandmed	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Rahvusvahelises multidistsiplinaarses AR	23	27
Rahvusvahelises valdkonnapõhises AR	31	46
Rahvuslikus multidistsiplinaarses AR	22	11
Rahvuslikus valdkonnapõhises AR	35	20
Keskuses institutsionaalses AR	34	30
Valdkonnapõhises institutsionaalses AR	45	46
Muu	11	11

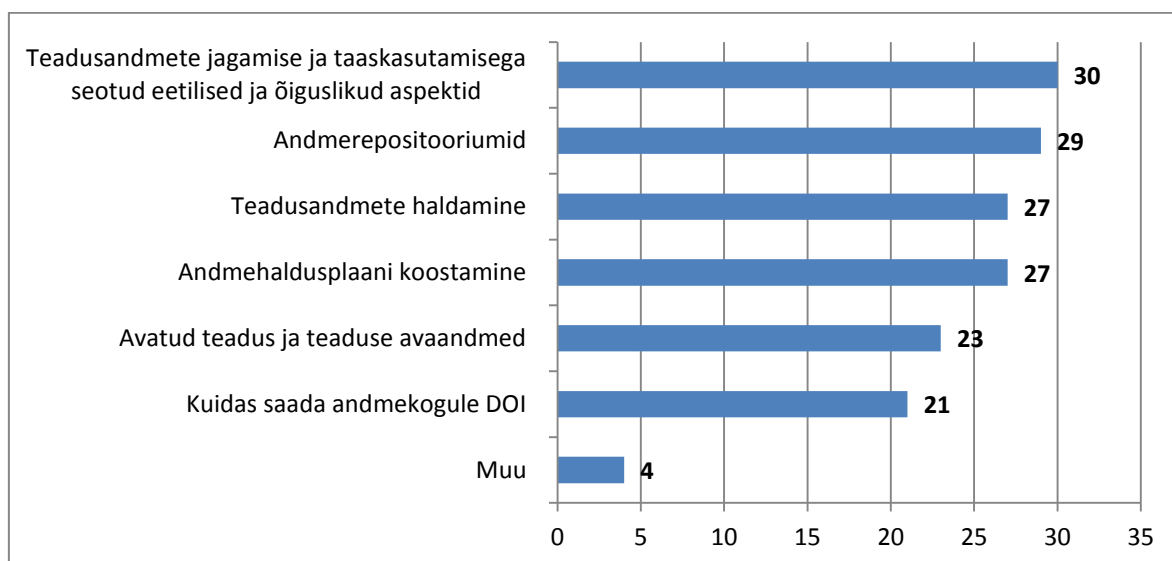
\*\*\* $p < 0,000$ , AR-andmerepositooriumis

Vastuseks küsimusele, kas oled kunagi täitnud rahastajate poolt nõutavat andmehaldusplaani, võis valida ühe vastusevariandi. Küsimusele vastas 61% kõigist vastanutest. Andmehaldusplaani ei ole midagi kuulnud ligi pooled (49%) vastanutest (joonis 18). 45% vastanutest ei ole andmehaldusplaani kunagi täitnud. Andmehaldusplaani on täitnud 6% vastanutest.



**Joonis 18. Kas olete täitnud andmehaldusplaani? (%)**

30% vastanutest sooviks rohkem infot teadusandmete jagamise ja taaskasutamisega seotud eetiliste ja õiguslike aspektide kohta (joonis 19). Lisainformatsioon andmerepositooriumide kohta huvitab 29% vastanutest.



**Joonis 19. Milliste teemade kohta sooviks rohkem infot? (%)**

Kommentaaries täpsustati: „On prognoositud, et varsti tuleb suurematelt teleksoopidelt suurusjärgus terabait andmeid päevas, mille haldamine, isegi salvestamine, on raske. Oleks huvitatud, mis on olemasolevad võimalused niisuguse andmemahuga toimetulekuks“. Samuti huvitavad: „Arengud globaalse "Linked Data" ja "Web of Data" suundadel“ ning „Kuidas pdf andmeid digitaliseerida ja viia probleemivabalt Excelisse“. Sotsiaalteaduste valdkonna vastajad on kõige enam huvitatud lisainfost andmehaldusplaani koostamise kohta (tabel 16). Loodus- ja täppisteaduste valdkonna vastajaid huvitab enim teadusandmete jagamise ja taaskasutamisega seotud eetilised ja õiguslikud aspektid.

**Tabel 16. Milliste teemade kohta sooviks rohkem infot? Valdcondlik jaotus**

Milline lisainfo pakub huvi?	Sotsiaalteadused (%)	Loodus- ja täppisteadused (%)
Andmehaldusplaani koostamine	55	40
Teadusandmete haldamine	49	41
Eetilised ja õiguslikud aspektid	51	48
Andmerepositooriumid	48	46
Avatud teadus ja teaduse avaandmed	46	34
Kuidas saada andmekogule DOI	39	33

\*\*\* $p < 0,000$



## 5 ARUTELU

Uuringu eesmärgiks oli välja selgitada ja kirjeldada, kuidas Tartu Ülikooli loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna teadlaste ja doktorantide seas teadusandmeid kogutakse, säilitatakse, jagatakse ja taaskasutatakse.

Enamus vastajatest tegeles eelmisel aastal ühe uurimisprojektiga ning kogus andmeid uurimisprojekti liikmena konkreetse uurimistöö käigus.

Teadusandmed, mida kogutakse on eritüübilised. Kogutakse kõiki küsimustikus esitatud andmetüüpe, kuid enam levinud on andmete kogumine tabeli ja teksti kujul. Sarnased tulemused saadi ka Oxfordi ülikoolis (Wilson, 2012) ja Exeteri ülikoolis (Team, 2012) läbi viidud uuringutes. Kolmandik vastanutest kogub andmeid ka andmebaasina. Siinkohal tuleb märkida, et vastajal võis tekkida raskusi andmebaasi ja tabeli vahel valiku tegemisel. Uurija pidas silmas eelkõige andmebaasirakendusi nagu nt MySQL, MS Access. Võimalik, et osa vastajaid valis tabeli asemel andmebaasid ja vastupidi. Samas ei saa välistada, et kasutusel ongi mõlemad variandid.

Loodus- ja täppisteaduste valdkonnas on tähelepanuväärne pildiliste andmete suur osakaal, mis võib olla seotud ka Tartu Ülikooli osalusega Eesti eElurikkuse (<http://elurikkus.ut.ee/>) andmekogu arendamisel. Pooled andmetest kogutakse instrumentidega, mis on samuti loomulik arvestades Tartu Ülikooli varustatust kõrgtehnoloogiliste uurimislaboritega.

Sotsiaalteaduste valdkonnas kogutakse ligi kolmandik andmeid helifailidena. Arvata võib, et tegemist on peamiselt intervjuude salvestistega.

Tähelepanuväärselt on ligi veerand sotsiaalteaduste valdkonna vastanute andmetest instrumendiga seotud andmed. Siinkohal oleks edaspidi huvitav teada, milliseid instrumente silmas peetakse. Võimalik, et mõeldi näiteks diktofoni või videokaamerat.

Kolm neljandikku vastanutest ei ole pidanud teadusandmete kaoga kokku puutuma. Nagu tulemustest näha kombineeritakse erinevaid andmete varundamise taktikaid. Andmeid eelistatakse hoida uurimisprojekti vältel kõige enam töö- või isiklikus arvutis. Umbes kolmandik vastajaid kasutab mälupulka, väiksem arv vastajaid pilveserverit ning arvuti välist kõvaketast. Oluline on, et ei kasutataks ainult ühte neist variantidest. Leopando järgi on rusikareeglid andmete hoidmiseks, et kolm koopiat on erinevatel andmekandjatel või võrguketastel, kahes eri formaadis ning üks neist tegevuskohaväline (Leopando, 2013). Üldine soovitus on mälupulka kasutada vaid andmete transportimiseks mitte (Briney, 2015: 118) Samuti ei tohiks pilveserver olla ainuke andmete säilitamise lahendus

(Leopando, 2013). Tulemustest ilmneb, et peamiselt inimliku hooletuse või tehnika tõrkumise tõttu on ligikaudu veerand vastanutest siiski ka oma andmetest ilma jäänud.

Tartu Ülikooli IT osakonna poolt hallatavat võrguketast kasutab andmete hoidmiseks alla veerandi vastanutest. Tegemist on institutsionaalse võrgukettaga, mille usaldusväärsus on tagatud. Võimalik, et selle vähene kasutamine viitab sellele, et uuringus osalenud ei ole antud võimalusest piisavalt informeeritud või teadlikud riskidest, mis kaasnevad paljude teiste andmete hoidmise viisidega.

Peamiselt nähakse vastutust andmete säilitamise eest, nii uuringuprojekti vältel kui ka pärast seda iseendal ning kõige meelsamini säilitatakse andmeid määramatult. Samas tundub, et pikaajaline säilitamise strateegia ei ole läbi mõeldud. Tõenäoliselt lihtsalt ei kustutata uuringute andmeid teadlikult oma arvutist. Andmepositooriumis hoiab oma andmeid alla 10% vastanutest. Töö- ja isikliku arvuti kõrval on kasutusel kolmandal kohal väline kõvaketas. Võib arvata, et andmete pikaajalise säilitamise puuduste põhjusteks on oskuste ja info puudus andmete arhiveerimise osas. Sellele viitab ka asjaolu, et ligi 30% vastanutest on huvitatud lisainformatsioonist andmepositooriumide kohta. Samuti on lisakommentaaries toodud välja tugiteenuste vajalikkus. Siiski võib öelda, et valdkondlikult kasutavad loodus- ja täppisteaduste valdkonna vastanud rohkem nii Tartu Ülikooli võrguketast kui ka andmepositooriume. Viimane võib olla seotud sellega, et loodus- ja täppisteaduste valdkonna teadusajakirjad nõuavad üha enam artiklite avaldamisel ka algandmete linkimist andmepositooriumist.

Repositooriumid nagu näiteks ka vastanute poolt viidatud Dryad (<http://datadryad.org/>) registreerivad teadusandmete metaandmed ja omistavad DOI-numbri ehk digitaalobjekti identifikaatori (*Digital Object Identifier*) ning muudavad seeläbi teadusandmed maailmas nähtavaks. Uuringus osalenud töid näidetena välja ka Eesti teadlaste poolt hallatavad repsitooriumid QsardDB (<http://qsardb.org/repository/>) ja PlutoF (<https://plutof.ut.ee/>), mis samuti registreerivad Tartu Ülikooli raamatukogu vahendusel teadusandmete metaandmed ja omistavad DataCite DOI-numbrid (<https://www.datacite.org/>)

Uuringu andmetest lähtuvalt, töötavad loodus- ja täppisteadlased rohkem uurimisgrupis. Teadusandmete säilitamise ja kättesaadavaks tegemise teema muutub kogu maailmas üha tähtsamaks, seega võib arvata, et eelkõige rahvusvahelistes uurimisgruppides töötades, on rohkem kokku puutunud ka andmete arhiveerimispraktikatega.

Märkimisväärne hulk, ligi veerand, andmetest säilitatakse paberil. Kusjuures uurimisprojekti ajal on isegi pea kolmandik andmetest paberkandjal. Kuigi loodus- ja täppisteaduste valdkonna vastanud kasutavad rohkem andmepositooriume, siis

paradoksaalselt ilmneb valdkondliku erinevusena, et loodus- ja täppisteaduste valdkonna vastanud kasutavad andmete säilitamiseks paberkandjat rohkem kui sotsiaalteaduste valdkonna vastanud. Võib oletada, et tegemist on laborimärkmete või välitööde päevikutega. Igal juhul oleks edaspidi huvitav teada, mis põhjustel ning milliste andmete säilitamiseks kasutatakse paberkandjat. Võimalik, et osaliselt on selleks ka harjumus või turvatunde puudumine digitaalsete keskkondade suhtes. Sellele võiks viidata vastanu kommentaar, et andmeid hoitakse „paberitel juhendaja lukustatud kapis“. Huvitav oleks uurida, et kas paberkandjatel olevaid andmeid on tulevikus plaanis ka digiteerida.

Kui vastajatele anti valida, millises andmerekpositooriumis nad oma andmeid arhiveeriksid, siis eelistati valdkonnapõhiseid repositooriume nii rahvusvahelisi kui ka rahvuslikku. See on mõistetu, sest tõenäosus, et sarnaste huvidega kolleegid andmed üles leiavad ja neid tsiteerivad on suurem, kui andmeid arhiveeritakse valdkondlikes repositooriumides (Briney, 2015: 157). Loodus- ja täppisteaduste valdkonna vastanud eelistaksid rahvusvahelist valdkonnapõhist ja valdkonnapõhist institutsionaalset andmerekpositooriumi. Sotsiaalteaduste valdkonna vastanud valdkonnapõhist institutsionaalset ning rahvuslikku valdkonnapõhist andmerekpositooriumi. Rahvuslikke andmerekpositooriume valiksid pigem sotsiaalteaduste valdkonna vastanud. Võib oletada, et selle põhjuseks on, et paljud sotsiaalteaduste valdkonna uuringud on valdavalt Eestikesksed.

Lisaks sellele, et andmeid soovitakse säilitada määramatult, arwab enamus (81%) vastanutest, et nende kogutud andmed võiksid, kas kindlasti või siis mõnikord, ka teistele huvi pakkuda. Nagu üks vastanutest kommenteerib: „Ma ei tee ju tööd tegemise pärast. Ikka loodan, et kunagi oleks sellest kasu.“ Samas lisavad vähem kui pooled vastanutest oma andmetele kirjeldused ja metaandmed, mis on andmete leidmiseks ning sisu mõistmiseks olulised.

Teadusandmete jagamise suhtub suurem osa vastanutest pooldavalt. Ligi 70% vastanutest on valmis oma andmeid teistega jagama, kuid seda tehakse enamasti siis, kui keegi on neist konkreetselt huvitatud ning andmed edastatakse peamiselt füüsiliste andmekandjate või e-posti kaudu. Pigem ollakse valmis andmeid jagama teadlaskonnaga kui kogu avalikkusega. Peamise kasuna nähakse eelkõige uusi koostöövõimalusi teiste teadlastega ning seeläbi oma teadusvaldkonna kiiremat arengut. Sarnase tõdemuseni jõudsid Fecher jt uurides teadusandmete jagamise motivaatoreid. (Fecher, Friesike, & Hebing, 2015:25) Umbes kolmandik vastanutest jagaksid andmeid meelsamini, kui sellega arvestataks ka akadeemilises hindamisprotsessis.

Sotsiaalteaduste valdkonna vastanute puhul leiavad üle poole vastanutest, et andmete jagamine vähendab liigset andmete kogumist. Samas loodus- ja täppisteaduste valdkonna vastanutest arvab sama, alla poole vastanutest. Võimalik, et põhjuseks on, et sotsiaalteaduste valdkonnas kasutatakse uuringutes näiteks Euroopa Sotsiaaluuringu (Tartu Ülikooli ühiskonnateaduste..., 2016) ja Eesti Statistika (<http://www.stat.ee/>) andmestikke, mis juba ongi avalikud. Samuti tehakse Eesti Vabariigi ministeeriumide tellimusel uuringuid riiklikest registritest kättesaadavate andmete põhjal. Põhjusi, miks loodus- ja täppisteaduste valdkonna vastanud antud väitega sama päri ei ole, tasuks edasi uurida.

Peamisteks takistuseks andmete jagamisel on soov neid edaspidi oma tulevases uurimistöös kasutada ning konkurents ja võistlus teaduses. Sotsiaalteaduste valdkonna vastanud peavad samavõrd suureks takistuseks ohtu privaatsusele, mis on mõistetav, sest paljud uuringud on seotud tundlike andmete kogumisega. Varasemates uuringutes on leitud, et sotsiaalteaduste valdkonnas jagatakse andmeid teiste valdkondadega (v. a meditsiin) võrreldes vähem meelsamini. (Tenopir et al., 2011:12) Sarnaselt ilmnes ka käesolevast uuringust, et sotsiaalteaduste valdkonnas jagatakse andmeid eelkõige koostööpartneritega või inimestega, kellel on nende vastu konkreetne huvi, mitte avalikkusega.

Oma uuringus leidsid Tenopir jt, vanuselise erinevusena, et üle 50-aastased teadlased jagavad oma andmeid ilma piiranguteta meelsamini, kui kuni 50-aastased teadlased (Tenopir et al., 2011:15) Käesolevas uuringus vanusest lähtuvaid erinevusi ei ilmnenud.

Samuti ei ilmnenud vastajate vahel akadeemilisest positsioonist sõltuvaid erinevusi.

Teadusrahastajate poolt nõutava andmehaldusplaani koostamisega on kokku puutunud alla 10% vastanutest. Pooled uuringus osalenutest ei ole sellest teadlikud, mis on ka arusaadav, sest Eestis veel hetkel andmehaldusplaanide granditaotluste juures ei nõuta ning vastav praktika puudub.

Tundub, et teadusandmete haldamist, säilitamist ja kättesaadavaks tegemise teemad pakuvad uuringus osalenutele ka edaspidi huvi. Lisainfot soovitakse eelkõige andmete jagamise ja taaskasutamisega seotud eetiliste ja õiguslike aspektide kohta. Samuti andmepositooriumide, andmehaldusplaani koostamise ja teadusandmete haldamise kohta.

Kokkuvõtvalt võib öelda, et teadusandmete väärtust hinnatakse kõrgelt ja nähakse nende kasu ka tulevikus. Andmeid ollakse valmis teistega jagama, kuid praegu on see tegevus veel juhuslikku laadi või põhineb uurimisgrupisisestel praktikatel. Päris täpselt ei teata, kuidas andmeid tuleks hallata ning arhiveerida, tuntakse puudust juhenditest ja

tugiteenustest. Andmerepositooriumidega ollakse kokku puutunud vähe. Samas pooldatakse avatud teaduse mõtteviisi. Soovitav oleks kehtestada juhendumiseks konkreetsed riiklikud ja institutsionaalsed poliitikad, millega Eesti Teadusagentuuri juhtimisel ka tegeletakse, välja töötada andmehaldusteemalised juhendid ja õppematerjalid ning tegeleda teavitustööga, kuidas teadusandmeid kättesaadavaks teha.

## KOKKUVÕTE

Tänu infotehnoloogia kiirele arengule on tänapäeva teadus muutunud andmemahukaks, samas on teadusandmete kogumine ja jagamine muutunud varasemast kiiremaks ja lihtsamaks. Üha kasvavad andmehulgad toovad kaasa vajaduse andmete eest hoolitseda nii, et need oleksid ka hiljem kasutatavad.

Ülemaailmsed globaliseerumis- ja integreerumisprotsessid on rajanud tee avatud teaduse paradigma väljakujunemisele, mis annab teadlaste töö tulemused, nii teadusartiklid kui ka nende aluseks olevad teadusandmed, tasuta kogu ühiskonna käsutusse.

Avatud teaduse mõtteviis on viinud avatud teaduse poliitikate väljatöötamiseni kogu maailmas, Euroopa Liidus, paljudes riikides ja institutsioonides. Olulise initsiatiivi on selles haaranud teaduse rahastajad, kes soovivad vältida sarnaste uuringute korduvat toetamist ja ootavad, et nende rahastatud projektid tooksid maksimaalset kasu ühiskonnale ka pikas perspektiivis.

Kõige olulisem roll teadusmaailmas on teadlastel endil. Teadlase kõige suurem varandus on teaduspublikatsioonid ja nende aluseks olevad andmed, mis määravad kogu akadeemilise karjääri. Seetõttu ei saa eeldada, et kõik uurijad oma töö tulemusi meelsasti jagavad, ehkki jagamine soodustab valdkondadeülest ja rahvusvahelist koostööd, aitab kaasa uutele tõlgendustele ja avastustele ning toetab õppetööd. Teadustulemuste kvaliteeti saab hinnata vaid siis, kui alusandmed on leitavad, mõistetavad ja taaskasutatavad kas kordusuuringuks või uute hüpoteeside aluseks.

Praegu on teadlase isiklik ja silmnähtav kasu oma andmete jagamisest küllaltki väike, pigem nähakse andmete säilitamise ja jagamisega kaasnevat aja- ja rahakulu.

Eesti Teadusagentuuri eestvedamisel on kokku kutsutud avatud teaduse ekspertkomisjon, kelle ülesanne oli välja töötada ettepanekud avatud teaduse riikliku poliitika põhimõtete kujundamiseks. 2016. aasta veebruaris avaldati dokumendi „Avatud teadus Eestis“ tööversioon avalikkusele kommenteerimiseks.

Selle tõttu oli magistritöö raames otstarbekas uurida TÜ teadlaste suhtumist avatud teaduse mõtteviisi ning sellega seotud praktikaid, et teada saada, milline on avatud teaduse poliitika koostamiseks vajalik stardipositsioon.

Magistritöös uuriti teadusandmete kogumise, säilitamise, kättesaadavaks tegemise ja taaskasutamise praktikaid Tartu Ülikooli loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna teadlaste ja doktorantide seas.

Uurimismeetodiks oli küsitlusuuring. Selle tarbeks koostati ja viidi läbi veebiküsitlus LimeSurvey keskkonnas, millele vastas 391 teadlast ja doktoranti (22% potentsiaalsest valimist). Statistiliseks andmetötluseks kasutati SPSS ja Microsoft Excel tarkvara.

Küsimustik oli jaotatud viide plokki:

- teadusandmete kogumine;
- teadusandmete arhiveerimine;
- teadusandmete kättesaadavus, nähtavus ja taaskasutatavus;
- infrastruktuur ja teenused;
- andmed vastaja kohta.

Info selle kohta, kuidas ja milliseid andmeid kogutakse, paneb aluse tulemuste analüüsile ja teadusandmete haldamise meetoditele. Umbes 50% andmetest koguti tabeli või tekstina, kusjuures loodus- ja täppisteaduste valdkonnas tabeli kujul 85% ning sotsiaalteaduste valdkonnas tekstina 86%.

Huvitav on märkida, et juba praegu kasutatakse oma varem kogutud andmeid (52%) või teiste poolt kogutud andmeid (46%), mis ei ole küll avalikud, kuid see praktika räägib avatud teaduse poolt.

Selgus, et loodusteadlased töötavad valdavalt uurimisgrupis, sotsiaalteadlased võrdsel määral individuaalselt ja uurimisgrupis, kuid kõik vastutavad oma teadusandmete eest ise (72%) või on vastutus jagatud.

Projekti vältel hoitakse oma andmeid valdavalt tööarvutis ja/või isiklikus arvutis, järgnevad võrdselt paberkandjal ja mälupulgal säilitatavad andmed. Tunduvalt turvalisemad valikud, pilveserverid ja TÜ võrgukettad on vähem kasutusel.

Umbes veerand vastajatest on kokku puutunud andmekaoga, kuid sellele vaatamata ei ole andmete arhiveerimine eriti läbi mõeldud. Üldiselt säilitatakse andmeid pärast uurimisprojekti lõppu sealsamas, kus need juba paiknevad, kas siis isiklikus või tööarvutis või paberkandjal. Siiski on kasutusel ka välised kõvakettad, TÜ võrgukettad ning pilveserverid. Andmerepositooriumis, mis on pikas perspektiivis kõige turvalisem säilituskoht, hoiustab oma andmeid 12-16% teadlastest. Kõige vähem osati vastata küsimusele, kui kaua andmeid säilitatakse. Vaid 18% vastanutest oskas seda aega määratleda.

Pooled vastanutest arvasid, et nende andmed võiksid tulevikus teistele huvi pakkuda, 31% olid kahtleval seisukohal.

Andmete jagamine eeldab nende varustamist dokumentide ja metaandmetega. Selles osas oli praktika ja teadmised puudulikud, kirjeldavaid andmeid kasutab alla poole vastanutest. Üldiselt toetatakse avatud teaduse põhimõtteid ja usutakse teadusandmete jagamise positiivsesse mõjusse. Kõige enam nähakse uusi koostöövõimalusi ja oma teadusala arengut, aga ka võimalust teadustöö kvaliteeti hinnata ning tsiteeritavust ja teadustöö nähtavust suurendada.

Kõige rohkem motiveeriks teadlasi oma andmeid jagama tsiteeringute suurenemine ning selle mõju akadeemilisele hindamisprotsessile.

Selleks, et andmeid oleks võimalik säilitada ja jagada, tuleb nende haldamisega tegelda kogu uurimistöö vältel, mitte tagantjärele. Abistavaks vahendiks on seejuures andmehaldusplaan, mis ei ole veel teed praktikasse leidnud, kuid tõenäoliselt muutub kohustuslikuks elemendiks uurimistoetuse taotlemise protsessis.

Küsitlusest selgus, et kui teadlased peaksid tegelema andmehaldusega kogu andmete elutsükli jooksul, vajavad nad koolitust ja tugiteenuseid, mis seda protsessi toetaksid.

Tartu Ülikooli raamatukogu pakub mitmeid andmehalduseks vajalikke koolitusi ja teenuseid juba 2015. aastast. Uurimistööst saadud tulemused andmehalduse praktikate ja ootuste kohta võimaldavad jätkuvat teenuste väljatöötamist teadlaskonna toetuseks. Valdkondadevaheliste erinevuste teadvustamine võimaldab erialapõhist lähenemist.

Olles avatud teaduse eestkõnelejaks ülikoolis, annab raamatukogu oma panuse TÜ teadlaste mõjukuse suurendamiseks maailmas.



## SUMMARY

The aim of the present Master's thesis is to find out whether and how the University of Tartu researchers and doctoral students in the fields of natural and exact sciences and social sciences collect, preserve, share and re-use research data and what kind of additional information on the subject is still needed.

A web-based survey was carried out. 391 persons, comprising 22% of the whole sample, answered to the survey questions. The results were analysed by using quantitative methods. The answers revealed that less than 10% of the respondents had had any experience in building the data management plans required by the funders of research. Half of the respondents were even not aware of the possibility, which can be understood because at present, submitting of data management plans together with grant applications is not yet required in Estonia and there is no common practice for that.

The main obstacles in sharing research data are the wish to continue using them in one's own further research, and competition and rivalry in science.

In general, it can be said that research data are highly valued and seen to be of benefit in the future as well. People are ready to share their data, but at present the sharing is still rather sporadic or based on internal practices of the research group. Researchers are not sure about how to manage and archive their data, guidelines and supporting services are sorely needed. Only few respondents had had any experience in using data repositories. However, the idea of open science was welcomed by many respondents. It was also pointed out that concrete national and institutional policies regarding these issues and specific data management manuals and learning materials should be drawn and developed, and much attention should be paid to informing the scientific community about the ways of making research data accessible.

Key words: research data, research data management, data repositories, open science

## KASUTATUD KIRJANDUS

- Anglada, L., González Copeiro, C., Rico, P., Abadal, E., Bernal, I., Besson Ribas, C., ... others. (2014). *Recommendations for the Implementation of Article 37 of the Spanish Science, Technology and Innovation Act: Open Access Dissemination*. Vaadatud 05.05.2016 <http://repositori.uvic.cat:8888/handle/10854/3740>
- Arguillas, F., & Block, W. (2016). *R2: CISER's Data Quality Review and Reproduction of Results Service*. Poster, IDCC 2016. Vaadatud 07.05.2016 [http://www.dcc.ac.uk/sites/default/files/documents/IDCC16/54\\_Arguillas%20and%20Block%20-%20Poster%20IDCC%202016.pdf](http://www.dcc.ac.uk/sites/default/files/documents/IDCC16/54_Arguillas%20and%20Block%20-%20Poster%20IDCC%202016.pdf)
- Atwater, B. F., Carson, B., Griggs, G. B., Johnson, H. P., & Salmi, M. S. (2014). Rethinking turbidite paleoseismology along the Cascadia subduction zone. *Geology*, G35902.1. <http://doi.org/10.1130/G35902.1>
- Bauer, Bruno, Ferus, Andreas, Gorraiz, Juan, Gründhammer, Veronika, Gumpenberger, Christian, Maly, Nikolaus, ... Thorp, Jer. (2015). *Researchers and Their Data. Results of an Austrian Survey - Report 2015*. <http://dx.doi.org/10.5281/zenodo.34005>
- Bell, G. (2009). Foreword. T. Hey, S. Tansley, & K. Tolle (Toim), *The Fourth Paradigm* (1 edition). Microsoft Research. Vaadatud 15.03.2016 [http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th\\_paradigm\\_book\\_complete\\_lr.pdf](http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf)
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. (2003). Max Planck Institute for the History of Science. Vaadatud 01.04.2016 [http://www.fu-berlin.de/sites/open\\_access/Veranstaltungen/oa\\_berlin/poster/Berlin-Declaration\\_Simone-Rieger\\_MPIWG.pdf](http://www.fu-berlin.de/sites/open_access/Veranstaltungen/oa_berlin/poster/Berlin-Declaration_Simone-Rieger_MPIWG.pdf)
- Briney, K. (2015). *Data Management for Researchers: Organize, maintain and share your data for research success*. Exeter: Pelagic Publishing.
- Brussels Declaration on Open Access to Belgian Publicly Funded Research*. (2012). Vaadatud 02.05.2016 <http://oep.hypotheses.org/1085>
- Budapest Open Access Initiative*. (2002). Vaadatud 15.04.2016 <http://www.budapestopenaccessinitiative.org/background>
- CASRAI dictionary*. (2015). Vaadatud 09.05.2016 [http://dictionary.casrai.org/Research\\_data](http://dictionary.casrai.org/Research_data)

- Ceci, S. J. (1988). Scientists' Attitudes toward Data Sharing. *Science, Technology, & Human Values*, 13(1/2), 45–52.
- Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, & National Academy of Science. (2009). *Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age*. Washington, D.C.: National Academies Press. Vaadatud 11.04.2016 <http://www.nap.edu/catalog/12615>
- Committee on Issues in the Transborder Flow of Scientific Data, U.S. National Committee for CODATA, Commission on Physical Sciences, Mathematics, and Applications, & National Research Council. (1997). *Bits of Power: Issues in Global Access to Scientific Data*. Washington, D.C.: National Academies Press. Vaadatud 29.03.2016 <http://www.nap.edu/catalog/5504>
- Corti, L., Eynden, V. V. den, Bishop, L., & Woollard, M. (2014). *Managing and Sharing Research Data: A Guide to Good Practice* (1 edition). Los Angeles: SAGE Publications Ltd.
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2014). *Managing and Sharing Research Data: A Guide to Good Practice* (1 edition). Los Angeles: SAGE Publications Ltd.
- Cox, A. M., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship & Information Science*, 46(4), 299–316. <http://doi.org/10.1177/0961000613492542>
- Cybernetica AS. (2016). *Sharemind*. Vaadatud 09.05.2016 <http://cyber.ee/infoturve/sharemind/>
- Dempster, S. (2014). *Ireland: The Transition to Open Access - Ireland Case Study*. Vaadatud 04.05.2016 <http://www.pasteur4oa.eu/sites/pasteur4oa/files/resource/Ireland%20Case%20Study.pdf>
- EDINA. (2015). *MANTRA. Research Data Management Training*. Vaadatud 07. 05.2016 <http://datalib.edina.ac.uk/mantra/>
- Eesti teadus- ja arendustegevuse ning innovatsiooni strateegia 2014-2020 „Teadmistepõhine Eesti”. (2014). Tartu: Haridus- ja Teadusministeerium. Vaadatud 22.04.2016 [https://www.hm.ee/sites/default/files/59705\\_teadmistepohine\\_eesti\\_est.pdf](https://www.hm.ee/sites/default/files/59705_teadmistepohine_eesti_est.pdf)

- Eesti Teadusagentuur. (2015). *Avatud teaduse ekspertkomisjon*. Vaadatud 08.05.2016  
<http://www.etag.ee/tegevused/teemad/avatud-teadus/avatud-teaduse-ekspertkomisjon/>
- Eesti Teadusagentuuri avatud teaduse ekspertkomisjon. (2016). *Avatud teadus Eestis. Eesti Teadusagentuuri avatud teaduse ekspertkomisjoni põhimõtted ja soovitused riikliku poliitika kujundamiseks*. Vaadatud 08.05.2016 [http://www.etag.ee/wp-content/uploads/2016/03/AT\\_soovitused\\_20160229.pdf](http://www.etag.ee/wp-content/uploads/2016/03/AT_soovitused_20160229.pdf)
- Euroopa Komisjon. (2012). *Euroopa Komisjoni soovitus, 17. juuli 2012, teadusinfo kättesaadavuse ja säilitamise kohta* (No. 2012/417/EL). Euroopa Liidu Teataja. Vaadatud 08.05.2016 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:ET:PDF>
- European Commission. (2016a). *European Open Science Cloud*. Vaadatud 12.04.2016  
<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>
- European Commission. (2016b). *Guidelines on Data Management in Horizon 2020*. European Commission. Vaadatud 12.04.2016  
[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)
- Expert Advisory Group on Data Access (EAGDA). (2014). *Establishing incentives and changing cultures to support data access*. MRC, ESRC, Cancer Research UK, Wellcome Trust. Vaadatud 31.03.2016 <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/EAGDA/WTP056496.htm>
- Fecher, B., Friesike, S., & Hebing, M. (2015). What Drives Academic Data Sharing? *PLOS ONE*, 10(2), e0118053. <http://doi.org/10.1371/journal.pone.0118053>
- Fitzgerald, A., Fitzgerald, B., & Pappalardo, K. (2009). The Future of Data Policy. T. Hey, S. Tansley, & K. Tolle (Toim), *The Fourth Paradigm: Data-Intensive Scientific Discovery* (1 edition). Microsoft Research. Salvestatud  
[http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th\\_paradigm\\_book\\_complete\\_lr.pdf](http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf)
- Government of the Republic Slovenia. (2015). *National Strategy of Open Access to Scientific Publications and Research Data in Slovenia 2015-2020*. Salvestatud  
[http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National\\_strategy\\_for\\_open\\_access.pdf](http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National_strategy_for_open_access.pdf)
- Gray, J. (2009). Jim Gray on eScience: A Transformed Scientific Method. T. Hey, S. Tansley, & K. Tolle (Toim), *The Fourth Paradigm: Data-Intensive Scientific*

- Discovery* (1 edition). Redmond, Washington: Microsoft Research. Vaadatud 15.03.2016 [http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th\\_paradigm\\_book\\_complete\\_lr.pdf](http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf)
- Gray, J., Szalay, A. S., Thakar, A. R., Stoughton, C., & others. (2002). Online scientific data curation, publication, and archiving. *Astronomical Telescopes and Instrumentation* (lk 103–107). International Society for Optics and Photonics. Vaadatud 18.04.2016 <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=874929>
- Hallik, M. (2014, oktoober 10). Kellele ja milleks on vaja avaandmeid? *Sirp*. Salvestatud <http://www.sirp.ee/s1-artiklid/c21-teadus/kellele-ja-milleks-on-vaja-avaandmeid/>
- Hallik, M., & Tarkpea, T. (2016). Open Science Policy - A Twisted Road and Clear Aims. A. Degkwitz (Toim), *Bibliothek der Zukunft. Zukunft der Bibliothek: Festschrift für Elmar Mittler: Amazon.de: Andreas Degkwitz: Bücher*. Berlin: Walter De Gruyter.
- Henty, M., & others. (2008). Dreaming of data: the library's role in supporting e-research and data management. Vaadatud 24.03.2016 <https://digitalcollections.anu.edu.au/handle/1885/47617>
- Herzog, D. L. (2015). *Data Literacy: A User's Guide*. Los Angeles: SAGE Publications, Inc.
- Hey, T., & Trefethen, A. (2003). The data deluge: an e-Science perspective. F. Berman, G. C. Fox, & A. J. G. Hey (Toim), *Grid computing: making the global infrastructure a reality*. New York: J. Wiley. Vaadatud 24.03.2016 <http://blog.sijinhe.com/wp-content/uploads/2011/09/Grid-Computing-Making-the-Global-Infrastructure-a-Reality.pdf>
- HRK German Rectors`Conference. (2014). *Management of research data – a key strategic challenge for university management. Recommendation by the 16th General Meeting of the HRK on 13 May 2014 in Frankfurt/Main*. Vaadatud 07.05.2016 [https://www.hrk.de/uploads/tx\\_szconvention/HRK\\_Empfehlung\\_Forschungsdaten\\_13052014\\_EN.pdf](https://www.hrk.de/uploads/tx_szconvention/HRK_Empfehlung_Forschungsdaten_13052014_EN.pdf)
- Isikuandmete kaitse seadus*. (2007). Riigi Teataja. Vaadatud 07.05.2016 <https://www.riigiteataja.ee/akt/12802623>
- Jones, S., Whyte, A., & Pryor, G. (2013). *How to Develop Research Data Management Services-a guide for HEIs*. Edinburgh: Digital Curation Centre. Vaadatud 15.04.2016 <http://www.dcc.ac.uk/resources/how-guide>

- Karlstrøm, N., & Aasheim, J. H. (2014). *Education, Research and Open Access in Norway*. CRISTin. Vaadatud 07.05.2016  
[http://www.pasteur4oa.eu/sites/pasteur4oa/files/resource/Norway%20Case%20Study\\_0.pdf](http://www.pasteur4oa.eu/sites/pasteur4oa/files/resource/Norway%20Case%20Study_0.pdf)
- Kasberger, S. (2013). *Open Science*. Vaadatud 09.05.2016  
<http://openscienceasap.org/open-science/>
- Kraker, P., Leony, D., Reinhardt, W., & Beham, G. (2011). The case for an open science in technology enhanced learning. *International Journal of Technology Enhanced Learning*, 3(6), 643–654.
- Leonelli, S., Spichtinger, D., & Prainsack, B. (2015). Sticks and carrots: encouraging open science at its source. *Geo: Geography and Environment*, 2(1), 12–16.  
<http://doi.org/10.1002/geo2.2>
- Leopando, J. (2013). *World Backup Day: The 3-2-1 Rule*. Vaadatud 04.05.2016  
<http://blog.trendmicro.com/trendlabs-security-intelligence/world-backup-day-the-3-2-1-rule/>
- Lyon, L. (2012). The Informatics Transform: Re-Engineering Libraries for the Data Decade. *International Journal of Digital Curation*, 7(1), 126–138.  
<http://doi.org/10.2218/ijdc.v7i1.220>
- Ministry of Education and Culture. (2014). *Open science and research leads to surprising discoveries and creative insights. Open science and research roadmap 2014–2017*. Finland. Vaadatud 08.05.2016  
[http://www.minedu.fi/OPM/Julkaisut/2014/Avoimen\\_tieteen\\_ja\\_tutkimuksen\\_tiekartta\\_2014\\_2017.html?lang=en](http://www.minedu.fi/OPM/Julkaisut/2014/Avoimen_tieteen_ja_tutkimuksen_tiekartta_2014_2017.html?lang=en)
- Ministry of Higher Education and Science. (2014). *Denmark's National Strategy for Open Access*. Vaadatud 08.05.2016 <http://ufm.dk/en/research-and-innovation/cooperation-between-research-and-innovation/open-science/open-access-to-research-publications/engelsk-version-national-strategy-for-open-access.pdf>
- Murray-Rust, P. (2008). Open Data in Science. *Serials Review*, 34(1), 52–64.  
<http://doi.org/10.1016/j.serrev.2008.01.001>
- Muuli, V. (2014). *Research Data in Estonia: collecting, storing, availability. Some findings from questionnaire*. Vaadatud 07.05.2016 <http://dspace.ut.ee/handle/10062/44052>

- OECD Principles and Guidelines for Access to Research Data from Public Funding*. (2007). OECD. Vaadatud 28.03.2016 <http://www.oecd.org/sti/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm>
- Old Weather: Whaling*. (2016). Vaadatud 31.03.2016 <https://whaling.oldweather.org/#/>
- OpenAire+*. (2016). Vaadatud 07.05.2016 <https://www.openaire.eu/faq#article-id-236>
- Overview of funders' data policies*. (2016). Digital Curation Centre. Salvestatud <http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>
- Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., ... Dierolf, U. (2013). Making Research Data Repositories Visible: The re3data.org Registry. *PLOS ONE*, 8(11), e78080. <http://doi.org/10.1371/journal.pone.0078080>
- Paterson, M., Lindsay, D., Monotti, A., & Chin, A. (2007). DART: a new missile in Australia's e-research strategy. *Online Information Review*, 31(2), 116–134. <http://doi.org/10.1108/14684520710747185>
- Piowar, H. A., & Chapman, W. (2008). *A Review of Journal Policies for Sharing Research Data*. Toronto: Proceedings ELPUB 2008 Conference on Electronic Publishing. Vaadatud 15.04.2016 [http://elpub.scix.net/data/works/att/001\\_elpub2008.content.pdf](http://elpub.scix.net/data/works/att/001_elpub2008.content.pdf)
- Piowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175. <http://doi.org/10.7717/peerj.175>
- PLOS. (2016). *Editorial and Publishing Policies*. Vaadatud 09.05.2016 <https://www.plos.org/editorial-publishing-policies>
- Poline, J.-B., Breeze, J. L., Ghosh, S. S., Gorgolewski, K., Halchenko, Y. O., Hanke, M., ... Kennedy, D. N. (2012). Data sharing in neuroimaging research. *Frontiers in Neuroinformatics*, 6, 9. <http://doi.org/10.3389/fninf.2012.00009>
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Altmetrics: A manifesto*. Vaadatud 08.05.2016 <http://altmetrics.org/manifesto/>
- Recode Project Consortium. (2014). *Policy recommendations for open access to research data*. Vaadatud 15.04.2016 [http://recodeproject.eu/wp-content/uploads/2015/01/recode\\_guideline\\_en\\_web\\_version\\_full\\_FINAL.pdf](http://recodeproject.eu/wp-content/uploads/2015/01/recode_guideline_en_web_version_full_FINAL.pdf)
- Roosimäe, K. (2016, aprill 12). Info doktorantide arvu kohta. Isiklik kirjavahetus.
- Savage, C. J., & Vickers, A. J. (2009). Empirical Study of Data Sharing by Authors Publishing in PLoS Journals. *PLOS ONE*, 4(9), e7078. <http://doi.org/10.1371/journal.pone.0007078>

- Schäfer, A., Pampel, H., Pfeifferberger, H., Dallmeier-Tiessen, S., Tissari, S., Darby, R., ... Wilson, M. (2011). *Baseline Report on Drivers and Barriers in Data Sharing*. Vaadatud 15.04.2016 [http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/11/ODE-WP3-DEL-0002-1\\_0\\_public\\_final.pdf](http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/11/ODE-WP3-DEL-0002-1_0_public_final.pdf)
- Smith, K. B., & Hatemi, P. K. (2012). OLS is AOK for ACE: A Regression-Based Approach to Synthesizing Political Science and Behavioral Genetics Models. *Political Behavior*, 35(2), 383–408. <http://doi.org/10.1007/s11109-012-9192-2>
- Suber, P. (2012). *Open access*. Cambridge, Massachusetts: MIT Press. Vaadatud 15.04.2016 [https://mitpress.mit.edu/sites/default/files/9780262517638\\_Open\\_Access\\_PDF\\_Version.pdf](https://mitpress.mit.edu/sites/default/files/9780262517638_Open_Access_PDF_Version.pdf)
- Summary of the Report of the Second International Strategy Meeting on Human Genome Sequencing (Bermuda, 27th February - 2nd March, 1997) as reported by HUGO.* (1997). Human Genome Project Information Archive 1990-2003. Vaadatud 12.04.2016 [http://web.ornl.gov/sci/techresources/Human\\_Genome/research/bermuda.shtml](http://web.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml)
- Swedish Research Council. (2015). *Proposal for National Guidelines for Open Access to Scientific Information*. Vaadatud 08.05.2016 <https://publikationer.vr.se/en/product/proposal-for-national-guidelines-for-open-access-to-scientific-information/>
- Zimmerman, A. S. (2008). New Knowledge from Old Data The Role of Standards in the Sharing and Reuse of Ecological Data. *Science, Technology & Human Values*, 33(5), 631–652. <http://doi.org/10.1177/0162243907306704>
- Tartu Ülikool. (2012). *Tartu Ülikooli akadeemilise personali ametijuhend*. Vaadatud 05.01.2016 <http://bioinfo.ut.ee/web/wp-content/uploads/2013/06/Tartu-%C3%9Clikooli-akadeemilise-personali-ametijuhend.pdf>
- Tartu Ülikooli kodulehekülg.* (2016). Vaadatud 09.05.2016 <http://www.ut.ee/et>
- Tartu Ülikooli raamatukogu arengukava 2016-2020.* (2015). Tartu Ülikooli raamatukogu. Vaadatud 07.05.2016 <https://utlib.ut.ee/sites/default/files/dokumendid/T%C3%9CR%20ARENGUKAVA%202016-2020.pdf>
- Tartu Ülikooli raamatukogu kodulehekülg. (2016). *Avatud teadus*. Vaadatud 07.05.2016 <https://utlib.ut.ee/avatud-teadus-open-science>



- Tartu Ülikooli ühiskonnateaduste instituut. (2016). *Euroopa Sotsiaaluuring*. 2016.  
Vaadatud 29.04.2016 <http://www.yti.ut.ee/et/euroopa-sotsiaaluuring>
- Team, O. E. P. (2012). *Summary Findings of the Open Exeter Data Asset Framework (DAF) Survey*. University of Exeter. Vaadatud 03.05.2016  
<https://ore.exeter.ac.uk/repository/handle/10036/3689>
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., ... Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLOS ONE*, 6(6), e21101. <http://doi.org/10.1371/journal.pone.0021101>
- The Antarctic Treaty. (1961). *Australian Treaty Series 1961 No 12*. Canberra: Australian Government Publishing Services. Vaadatud 12.04.2016  
<http://www.austlii.edu.au/cgi-bin/sinodisp/au/other/dfat/treaties/1961/12.html?query=antarctic>
- The High Level Expert Group on Scientific Data. (2010). *Riding the Wave: How Europe can gain from the rising tide of scientific data*. European Union. Vaadatud 12.04.2016  
[http://ec.europa.eu/information\\_society/newsroom/cf/itemlongdetail.cfm?item\\_id=6204](http://ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=6204)
- The Research Council of Lithuania adopted its Guidelines on Open Access*. (2016). Vilnius University Library. Vaadatud 08.05.2016 <http://www.mb.vu.lt/en/news/scholarly-communication/research-council-lithuania-adopted-its-guidelines-open-access>
- Tonta, Y., Doğan, G., Al, U., & Madran, O. (2015). *Open Access Policies of Research Funders: The Case Study of the Austrian Science Fund (FWF)*. Departement of Information Management of Hacettwpw University. Vaadatud 08.05.2016  
[http://www.fwf.ac.at/fileadmin/files/Dokumente/Ueber\\_den\\_FWF/Publikationen/FWF-Selbstevaluation/FWF-OA-Policy-Case-Study\\_Pasteur4OA.pdf](http://www.fwf.ac.at/fileadmin/files/Dokumente/Ueber_den_FWF/Publikationen/FWF-Selbstevaluation/FWF-OA-Policy-Case-Study_Pasteur4OA.pdf)
- Tüli, K. (2016, aprill 9). Info personalinumbrite kohta. Isiklik kirjavahetus.
- University of Minho. (2014). *Portugal Open Access Policy Landscape*. Vaadatud 08.05.2016  
<http://www.pasteur4oa.eu/sites/pasteur4oa/files/resource/Portugal%20Case%20Study.pdf>
- UT - LimeSurvey*. (2016). Vaadatud 03.01.2016  
<https://www.ut.ee/survey/admin/admin.php>
- Van de Sompel, H., & Lagoze, C. (2009). All Aboard: Toward a Machine-Friendly Scholarly Communication System. T. Hey, S. Tansley, & K. Tolle (Toim), *The*

- Fourth Paradigm: Data-Intensive Scientific Discovery* (1 edition). Microsoft Research. Vaadatud 15.03.2016 [http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th\\_paradigm\\_book\\_complete\\_lr.pdf](http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf)
- Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). *Managing and sharing data: best practice for researchers*. Colchester: UK Data Archive.
- Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., ... Rennison, D. J. (2014). The Availability of Research Data Declines Rapidly with Article Age. *Current Biology*, 24(1), 94–97.  
<http://doi.org/10.1016/j.cub.2013.11.014>
- Whyte, A. (2015). *Where to keep research data: DCC checklist for evaluating data repositories* (Kd v.1.1). Edinburgh: Digital Curation Centre. Vaadatud 10.04.2016 [www.dcc.ac.uk/resources/how-guides](http://www.dcc.ac.uk/resources/how-guides)
- Wilson, J. (2012). *University of Oxford Research Data Management Survey 2012 : The Results / DaMaRO*. Vaadatud 08.02.2016  
<https://blogs.it.ox.ac.uk/damaro/2013/01/03/university-of-oxford-research-data-management-survey-2012-the-results/>

## LISA 1

**Tartu Ülikooli loodus- ja täppisteaduste valdkonnas töötanud  
akadeemiliste töötajate arv seisuga 09.03.2016**

Instituut	välistöötajad	Eesti töötajad
Arvutiteaduste instituut	29	77
Füüsika instituut	13	117
Keemia instituut	4	95
Eesti Mereinstituut	0	49
Molekulaar- ja rakubioloogia instituut	2	59
Matemaatika ja statistika instituut	0	45
Ökoloogia ja maateaduse instituut	12	153
Tehnoloogiainstituut	16	64
<b>Kokku</b>	<b>76</b>	<b>659</b>

**Tartu Ülikooli sotsiaalteaduste valdkonnas töötanud akadeemiliste  
töötajate arv seisuga 09.03.2016**

Instituut	välistöötajad	Eesti töötajad
Haridusteaduste instituut	2	56
Johan Skytte poliitikauuringute instituut	6	20
Majandusteaduskond	5	53
Narva kolledž	2	18
Õigusteaduskond	4	78
Pärnu kolledž	2	29
Psühholoogia instituut	0	28
Ühiskonnateaduste instituut	0	66
<b>Kokku</b>	<b>21</b>	<b>348</b>

**Tartu Ülikooli sotsiaalteaduste valdkonnas õppinud doktoriõppe  
üliõpilaste arv seisuga 07.03.2016**

Instituut	välisüliõpilane	Eesti üliõpilane
Haridusteaduste instituut	1	30
Johan Skytte poliitikauuringute instituut	7	8
Majandusteaduskond	9	53
Psühholoogia instituut	3	21
Õigusteaduskond	3	79
Ühiskonnateaduste instituut	2	40
Narva kolledž	0	0
Pärnu kolledž	0	0
<b>Kokku</b>	<b>25</b>	<b>231</b>

**Tartu Ülikooli loodus- ja täppisteaduste valdkonnas õppinud doktoriõppe  
üliõpilaste arv seisuga 07.03.2016**

<b>Instituut</b>	<b>välisüliõpilane</b>	<b>Eesti üliõpilane</b>
Arvutiteaduse instituut	20	36
Eesti mereinstituut		25
Füüsika instituut	9	73
Keemia instituut	5	65
Matemaatika ja statistika instituut	2	28
Molekulaar- ja rakubioloogia instituut	4	74
Tehnoloogiainstituut	8	44
Ökoloogia ja maateaduste instituut	7	167
<b>Kokku</b>	<b>55</b>	<b>512</b>

## LISA 2

### Sotsiaalteaduste valdkonna vastajad, kes määratlesid oma instituudi

Instituut	Vastajate arv	%
Haridusteaduste instituut	12	19
Johan Skytte poliitikauuringute instituut	1	2
Majandusteaduskond	11	18
Narva kolledž	3	5
Õigusteaduskond	2	3
Pärnu kolledž	4	6
Psühholoogia instituut	10	16
Ühiskonnateaduste instituut	19	31
<b>Kokku</b>	<b>62</b>	<b>100</b>

### Loodus- ja täppisteaduste valdkonna vastajad, kes määratlesid oma instituudi

Instituut	Vastanute arv	%
Arvutiteaduste instituut	6	3,5
Füüsika instituut	29	17,5
Keemia instituut	28	17
Eesti Mereinstituut	13	8
Molekulaar- ja rakubioloogia instituut	19	11,5
Matemaatika ja statistika instituut	6	3,5
Ökoloogia ja maateaduse instituut	48	29
Tehnoloogiainstituut	16	10
<b>Kokku</b>	<b>165</b>	<b>100</b>

## LISA 3

### Küsimustik

#### **Teadusandmete haldamise praktikad Tartu Ülikooli loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna teadlaste ja doktorantide seas**

Uuringu tulemused aitavad paremini mõista, kuidas teadlased ja doktorandid teadusandmeid koguvad, arhiveerivad, jagavad ja taaskasutavad. See teadmine aitab raamatukogul välja arendada teadlaskonna vajadustest lähtuvad teadusandmete haldamist toetavad teenused. Ühtlasi valmib uuringu tulemustest magistritöö.

Ankeedile vastamine võtab aega maksimaalselt 15 minutit.

Küsitlus on eesti keeles ning vastamine on anonüümne.

Kogutud andmeid kasutatakse vaid üldistatud kujul.

Küsimuste korral kirjutage või helistage Tartu Ülikooli raamatukogu teadusandmete spetsialistile Lilian Neerutile, 737 5729, lilian.neerut@ut.ee

### **TEADUSANDMETE KOGUMINE**

**Lisainfo:** Teadusandmed on digitaalandmed, mis on kogutud või loodud uurimistöö käigus, et neid analüüsida ning saada uusi originaalseid teadustulemusi.

Kommenteerimise võimalus on nupu "Muu" all

#### **Millist tüüpi on algandmed, mida uurimistöö käigus kogute?**

Palun valige **kõik** mis sobib:

- ☐ Tekst
- ☐ Tabel
- ☐ Audio
- ☐ Video
- ☐ Graafika/Pildid
- ☐ Andmebaasid
- ☐ Instrumendiga seotud andmed
- ☐ Mudelid
- ☐ Ruumilised andmed
- ☐ Tarkvara iseloomustavad andmed
- ☐ Muu

Soovi korral täpsustage

#### **Kuidas saate uurimistöö jaoks andmeid?**

Palun valige **kõik** mis sobib:

- ☐ Kogun andmeid konkreetse uurimistöö käigus
- ☐ Kasutan enda poolt varem kogutud andmeid
- ☐ Kasutan teiste poolt kogutud tasuta kättesaadavaid andmeid
- ☐ Ostan andmeid
- ☐ Muu

Soovi korral täpsustage

**Mitme uurimisprojektiga olite seotud eelmisel aastal?**

Palun valige **ainult üks** järgnevatest:

- ☐ 0
- ☐ 1
- ☐ 2-5
- ☐ 5 ja rohkem
- ☐ Muu

**Kas kogute teadusandmeid individuaalselt või uurimisgrupi liikmena?**

Palun valige **kõik** mis sobib:

- ☐ Individuaalselt
- ☐ Uurimisgrupi liikmena
- ☐ Muu

Soovi korral täpsustage

## **TEADUSANDMETE ARHIVEERIMINE**

**Lisainfo:** Andmepositooriumid on digitaalsed arhiivid teadusandmete ja nende metaandmete säilitamiseks ning kättesaadavaks tegemiseks.

Kommenteerimise võimalus on nupu "Muu" all

**Kus hoiate tavaliselt oma teadusandmeid uurimisprojekti vältel?**

Palun valige **kõik** mis sobib:

- ☐ Isiklikus arvutis
- ☐ Tööarvutis
- ☐ Pilveserveris (nt OneDrive, Dropbox)
- ☐ Tartu Ülikooli võrgukettal
- ☐ Arvuti välisel kõvakettal
- ☐ USB-l (mälupekk)
- ☐ CD/DVD-l
- ☐ Masinas või instrumendis, mis andmeid genereerib

- ☐ Paberkandjal (väljatrükid, laboriraamatud)
- ☐ Ei oska öelda
- ☐ Muu

Soovi korral täpsustage

**Kus hoiate tavaliselt oma teadusandmeid pärast uurimisprojekti lõppu?**

Palun valige **kõik** mis sobib:

- ☐ Andmerepositooriumis
- ☐ Isiklikus arvutis
- ☐ Töoarvutis
- ☐ Pilveserveris (nt OneDrive, Dropbox)
- ☐ Tartu Ülikooli võrgukettal
- ☐ Arvuti välisel kõvakettal
- ☐ USB-l (mälopulk)
- ☐ CD/DVD-l
- ☐ Masinas või instrumendis, mis andmeid genereerib
- ☐ Paberkandjal (väljatrükid, laboriraamatud)
- ☐ Ei oska öelda
- ☐ Muu

Soovi korral täpsustage

**Kui kaua plaanite säilitada teadusandmeid pärast uurimisprojekti lõppu?**

Palun valige **ainult üks** järgnevatest:

- ☐ Ei säilita
- ☐ 1 aasta
- ☐ 1-5 aastat
- ☐ 5-10 aastat
- ☐ Määramatult
- ☐ Ei oska öelda
- ☐ Muu

Soovi korral täpsustage

**Kes vastutab enamasti uurimisprojekti vältel teadusandmete säilitamise eest?**

Palun valige **ainult üks** järgnevatest:

- ☐ Mina ise
- ☐ Uurimisprojekti juht
- ☐ Tartu Ülikooli IT osakond
- ☐ Tartu Ülikooli raamatukogu



- ☐ Vastava kohustusega uurimisgrupi liige
- ☐ Teenusepakkuja väljaspoolt (**Soovi korral täpsustage**)
- ☐ Ei oska öelda
- ☐ Muu

Soovi korral täpsustage

**Kes vastutab enamasti pärast uurimisprojekti lõppu teadusandmete arhiveerimise eest?**

Palun valige **ainult üks** järgnevatest:

- ☐ Mina ise
- ☐ Uurimisprojekti juht
- ☐ Tartu Ülikooli IT osakond
- ☐ Tartu Ülikooli raamatukogu
- ☐ Vastava kohustusega uurimisgrupi liige
- ☐ Teenusepakkuja väljaspoolt (**Soovi korral täpsustage**)
- ☐ Ei oska öelda
- ☐ Muu

Soovi korral täpsustage

**Kas Teil on teadusandmeid kaduma läinud?**

Palun valige **ainult üks** järgnevatest:

- ☐ Jah
- ☐ Ei

Kommenteerige oma valikut siin:

## **TEADUSANDMETE KÄTTESAADAVUS, NÄHTAVUS JA TAASKASUTATAVUS**

**Lisainfo:** Avatud teaduse (Open Science) eesmärk on muuta teadusuuringud avatuks, globaalseks, loovaks ning koostööl põhinevaks, kasutades tänapäevaseid e-infrastruktuure.

Kommenteerimise võimalus on nupu "Muu" all

**Kas teadusandmed on varustatud metaandmete ja kirjeldustega nii, et need on hiljem arusaadavad ja kasutatavad Teile endale ja vajadusel ka kõigile teistele väljaspoolt uurimisgruppi?**

Palun valige **ainult üks** järgnevatest:

- ☐ Jah
- ☐ Ei
- ☐ Ei oska öelda
- ☐ Muu

Soovi korral täpsustage

**Kellele võimaldate juurdepääsu oma teadusandmetele, kui nende jagamine ei ole seotud piirangutega (nt konfidentsiaalsusnõue, ärisaladus)?**

Palun valige **kõik** mis sobib:

- ☐ Kõigile, kogu avalikkusele
- ☐ Kõigile teadlastele
- ☐ Kõigile oma eriala teadlastele
- ☐ Tartu Ülikooli liikmeskonnale
- ☐ Oma teaduskonna/instituudi/uurimiskeskuse liikmetele
- ☐ Oma konkreetsetele koostööpartneritele
- ☐ Huvitatud isikutele soovi korral
- ☐ Mitte kellelegi
- ☐ Muu

Soovi korral täpsustage

**Kuidas saavad teised kasutada Teie teadusandmeid?**

Palun valige **kõik** mis sobib:

- ☐ Andmerepositooriumi kaudu
- ☐ Andmed on lingitud repositooriumist lisamaterjalina teadusartikli juures
- ☐ Personaalse või institutsionaalse veebilehe kaudu
- ☐ Pilverakenduste kaudu (OneDrive, Dropbox, Google Drive, jne)
- ☐ Füüsiliste andmekandjate või e-posti kaudu
- ☐ Andmed ei ole kättesaadavad
- ☐ Ei oska öelda
- ☐ Muu

Võimalusel täpsustage repositooriumi nimi

Võimalusel täpsustage ajakirja nimi

Soovi korral täpsustage

**Kas Teie kogutud teadusandmeid võiksid tulevikus teistele teadlastele uurimistöö tegemiseks huvi pakkuda?**

Palun valige **ainult üks** järgnevatest:

- ☐ Jah
- ☐ Ei
- ☐ Mõnikord
- ☐ Ei oska öelda

Kommenteerige oma valikut siin:

**Mis motiveeriks Teid jagama oma teadusandmeid avatud teaduse põhimõttel (Open Science)?**

Palun valige **kõik** mis sobib:

- ☐ Tunnustus teadlaskonna poolt
- ☐ Akadeemilises hindamisprotsessis arvestatakse teadusandmetega kui olulise teadusliku väljundiga
- ☐ Andmete jagamine suurendab tsiteeritavust ning seeläbi kasvab minu teadustöö nähtavus ja mõjufaktor (ka ainult andmeid on võimalik tsiteerida)
- ☐ Uued kontaktid ja koostöövõimalused teiste teadlastega
- ☐ Rahalised boonused
- ☐ On olemas tugiteenused andmete jagamise protsessi lihtsustamiseks
- ☐ Riikliku ja institutsionaalse poliitika olemasolu
- ☐ Mitte miski
- ☐ Muu

Soovi korral täpsustage

**Miks on teaduse avaandmete jagamine Teie arvates kasulik?**

Palun valige **kõik** mis sobib:

- ☐ Aitab kaasa minu teadusvaldkonna arengule
- ☐ Aitab kaasa valdkondadevahelise teaduse arengule
- ☐ Suurendab koostöövõimalusi teiste teadlastega
- ☐ Vähendab liigset andmete kogumist
- ☐ Toetab avatud juurdepääsu teadmistele
- ☐ Kordusanalüüsid aitavad tõestada teadustulemusi
- ☐ Kaitseb andmete fabritseerimise ja võltsingute eest
- ☐ Andmete jagamine ja nende põhjal uuringute tegemine aitab õpetada tulevasi teadlasi
- ☐ Ei ole kasulik
- ☐ Muu

Soovi korral täpsustage

**Kas jagate meeledi oma teadusandmeid?**

Palun valige **ainult üks** järgnevatest:

- ☐ Jah
- ☐ Ei

### **Mis takistab Teid teadusandmeid jagamast?**

Palun valige **kõik** mis sobib:

- ☐ Minu teadusandmed on korrastamata
- ☐ Soovin neid veel oma tulevastes uurimistöödes kasutada
- ☐ Mul ei ole jagamiseks piisavalt teadmisi või tehnilisi oskusi
- ☐ Minu arvates ei pea teadusandmeid teistega jagama
- ☐ Ma ei teadnud, et teadusandmeid on võimalik jagada
- ☐ Ei ole kohta, kuhu teadusandmeid panna
- ☐ Minu teadusandmed ei paku teistele huvi ega ole neile kasulikud
- ☐ Privaatsuse rikkumise oht
- ☐ Teadusandmete väärkasutuse oht
- ☐ Valestitõlgendamise risk
- ☐ Muud legaalsed piirangud (copyright, patendiõigus, kaubamärgikaitse)
- ☐ Konkurents ja võistlus teaduses – võidab kiiremini publitseerija (põhimõte publish or perish)
- ☐ Teadusandmete ettevalmistamine jagamiseks on aja- ja töömahukas
- ☐ Teadusandmete säilitamiseks puuduvad standardid
- ☐ Kasutan väga haruldasi andmeformaate
- ☐ Teadusandmete võimalik kommertsiaalne kasutamine tulevikus
- ☐ Motivatsioonipuudus jagamiseks
- ☐ Muu

Soovi korral täpsustage

### **INFRASTRUKTUUR JA TEENUSED**

Kommenteerimise võimalus on nupu "Muu" all

### **Kus arhiveeriksid kõige meelsamini oma teadusandmeid?**

Palun valige **kõik** mis sobib:

- ☐ Rahvusvahelises multidistsiplinaarses andmerepositooriumis
- ☐ Rahvusvahelises valdkonnapõhises andmerepositooriumis
- ☐ Rahvuslikus multidistsiplinaarses andmerepositooriumis
- ☐ Rahvuslikus valdkonnapõhises andmerepositooriumis
- ☐ Keskuses institutsionaalses andmerepositooriumis (nt Tartu Ülikooli andmerepositooriumis datadoi.ut.ee)
- ☐ Valdkonnapõhises institutsionaalses andmerepositooriumis (nt instituudi või uurimisgrupi repositooriumis)
- ☐ Muu

Soovi korral täpsustage

**Kas oled kunagi täitnud rahastajate poolt nõutavat andmehaldusplaani (Data Management Plan)?**

Palun valige **ainult üks** järgnevatest:

- ☐Jah
- ☐Ei
- ☐Ei ole sellest midagi kuulnud

Kommenteerige oma valikut siin:

**Milliste teemade kohta sooviksite rohkem infot?**

Palun valige **kõik** mis sobib:

- ☐Andmehaldusplaani (Data Management Plan) koostamine
- ☐Teadusandmete haldamine
- ☐Teadusandmete jagamise ja taaskasutamisega seotud eetilised ja õiguslikud aspektid
- ☐Andmerepositooriumid
- ☐Avatud teadus ja teaduse avaandmed
- ☐Kuidas saada andmekogule DOI (Digital Object Identifier)
- ☐Muu

Soovi korral täpsustage

**Kui teate mõnda oma erialaspetsiifilist andmerepositooriumi, siis palun pange kirja.**

**Kas on veel midagi, mida soovite seoses teadusandmetega ära märkida?**

## **ANDMED VASTAJA KOHTA**

**Palun märkige millise Tartu Ülikooli valdkonnaga olete seotud?**

Palun valige **ainult üks** järgnevatest:

- ☐Loodus- ja täppisteaduste valdkond
- ☐Sotsiaalteaduste valdkond

Kommenteerige oma valikut siin:

**Palun märkige millise instituudiga olete seotud**

Palun valige **ainult üks** järgnevatest:

- ☐Arvutiteaduse instituut
- ☐Eesti Mereinstituut
- ☐Füüsika instituut
- ☐Keemia instituut
- ☐Matemaatika ja statistika instituut
- ☐Molekulaar- ja rakubioloogia instituut
- ☐Tehnoloogiainstituut

- ☐ Õkoloogia ja maateaduste instituut

**Palun täpsustage millise teaduskonna, instituudi või kolledžiga olete seotud**

Palun valige **ainult üks** järgnevatest:

- ☐ Õigusteaduskond
- ☐ Majandusteaduskond
- ☐ Ühiskonnateaduste instituut
- ☐ Haridusteaduste instituut
- ☐ Psühholoogia instituut
- ☐ Johan Skytte poliitikauuringute instituut
- ☐ Narva kolledž
- ☐ Pärnu kolledž

**Palun täpsustage millise õppetooliga olete seotud**

Palun valige **ainult üks** järgnevatest:

- ☐ Ajakirjanduse õppetool
- ☐ Infoteaduste õppetool
- ☐ Meediauuringute õppetool
- ☐ Sotsiaalpoliitika õppetool
- ☐ Sotsiaalse kommunikatsiooni õppetool
- ☐ Sotsiaaltöö õppetool
- ☐ Sotsioloogia õppetool

**Milline on Teie akadeemiline positsioon?**

Palun valige **ainult üks** järgnevatest:

- ☐ Professor
- ☐ Dotsent
- ☐ Lektor
- ☐ Assistent
- ☐ Juhtivateadur
- ☐ Vanemteadur
- ☐ Teadur
- ☐ Nooremteadur
- ☐ Doktorant

Kommenteerige oma valikut siin:

**Palun täpsustage mitmenda aasta doktorant olete**

Palun valige **ainult üks** järgnevatest:

- ☐ 1. aasta

- ☐ 2. aasta
- ☐ 3. aasta
- ☐ 4. aasta

**Teie vanus**

Kirjutage vastus siia:

**Suur tänu koostöö eest!**

**Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks**

Mina, Lilian Neerut,

*(autori nimi)*

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Teadusandmete haldamise ja jagamise praktikad Tartu Ülikooli loodus- ja täppisteaduste ning sotsiaalteaduste valdkonna teadlaste ja doktorantide seas“,  
*(lõputöö pealkiri)*

mille juhendajad on Kurmo Konsa (PhD) ja Mai Beilmann (MA)

*(juhendaja nimi)*

- 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
  - 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
  3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 10.05.2016